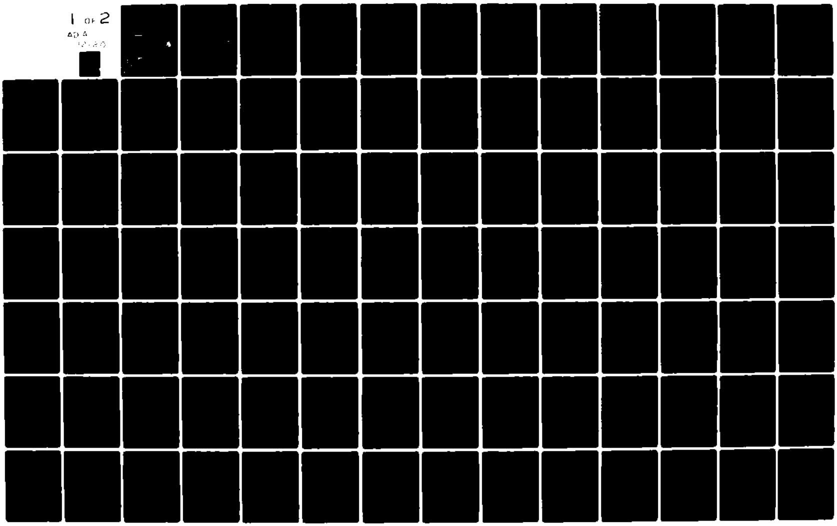
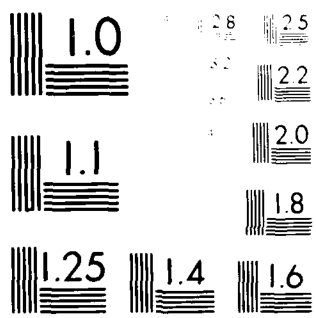
AD-A112 120   ILLINOIS UNIV AT URBANA COORDINATED SCIENCE LAB   F/G 12/2
A GENERAL APPROACH TO MINIMAX ROBUST FILTERING.(U)
DEC 81   S VERDU   N00014-79-C-0424
UNCLASSIFIED   R-933   ARO-17761.6-EL   NL

1 OF 2
AD A
112120





MICROCOPY RESOLUTION TEST CHART

ARO 17761.6-EL

(1)

REPORT R-933 DECEMBER, 1981 UILU-ENG 81-2264

COORDINATED SCIENCE LABORATORY

A GENERAL APPROACH TO MINIMAX ROBUST FILTERING

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A112120	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A GENERAL APPROACH TO MINIMAX ROBUST FILTERING		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-933; UILU-ENG 81-2264
7. AUTHOR(s) Sergio Verdú		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424; DAAG-81-K-0062
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program; U. S. Army Research Office		12. REPORT DATE December 1981
		13. NUMBER OF PAGES 99
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

minimax robust filtering; robust matched filters; minimax observers and regulators

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

General results systematizing the solution of minimax robust filtering problems are presented. Their application is investigated in the areas of matched filtering and of state estimation and control for linear time-varying stochastic systems. Further minimax filtering situations are studied for other problems in signal detection and estimation.

UILU-ENG-81-2264

A GENERAL APPROACH TO MINIMAX ROBUST FILTERING

by

Sergio Verdú

This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy and U.S. Air Force) under Contract N00014-79-C-0424 and in part by the U.S. Army Research Office under Contract DAAG-81-K-0062.

ABSTRACT

General results systematizing the solution of minimax robust filtering problems are presented. Their application is investigated in the areas of matched filtering and of state estimation and control for linear time-varying stochastic systems. Further minimax filtering situations are studied for other problems in signal detection and estimation.

ACKNOWLEDGMENTS

It is a pleasure for me to thank my advisor, Prof. Vincent Poor, for his continuous encouragement and support throughout this work. I am also grateful to Profs. D. P. Looze, M. Mintz and J. J. Uhl, Jr., and to G. Bilardi and K. S. Vastola for helpful discussions and comments, and to Ms. Cathy Cassells for her splendid typewriting of the sometimes awkward manuscript. Finally, I want to acknowledge the funding provided by the Joint Services Electronics Program (U.S. Army, U.S. Navy, U.S. Air Force) and the U.S. Army Research Office under Contracts N00014-79-C-0424 and DAAG-81-K-0062.

TABLE OF CONTENTS

Contents

1. INTRODUCTION

When some of the quantities involved in the mathematical model that describes a particular filtering situation are not completely known, another model incorporating the available partial knowledge must be adopted. This partial information can be, for example, in the form of a prior distribution function if considering the unknown quantity as a random variable, or in the form of an uncertainty set, any of the members of which can be the actual quantity. Then the objective could be, in the first case, to maximize the expected value of the original performance measure or, in the second case, to maximize the performance of the optimum operating point over the uncertainty set.

Focusing our attention in the case in which the unknown quantity is modeled by an uncertainty set, suppose that our aim is to synthesize a system that is somehow insensitive (*robust*) to the member of the uncertainty set that is actually present. One of the possible strategies to achieve this is to optimize the worst-case performance (*minimax*). In such case, for any other system there exists an operating point in the uncertainty set for which it performs worse than the minimax system. This can be considered as a game in which the designer tries to guarantee a maximum payoff assuming that his opponent (nature) will select the worst possible operating point for whatever system he uses.

Because the less uncertain the mathematical model is the better the resulting system will be, an alternative that the engineer may consider is that of an adaptive system; i.e., one consisting of a subsystem that narrows down the uncertainty by learning a more accurate model of the unknown system from the incoming observations. This is especially necessary when the uncertainties are so large that trivial and robust filtering perform similarly. Indeed the adaptive and minimax philosophies constitute the mainstreams of the design of uncertain systems. The mutual advantages and disadvantages of both approaches have been widely discussed. To name a few, adaptive systems are supposed to be more complex, to have worse performance for small sample size problems, and thus inferior dynamical response, and to be, commonly, ad-hoc solutions; minimax systems are criticized as being pessimistic (although often this is not the least advisable engineering approach), as losing performance in some nominal (and maybe most probable) model, and as being dependent on the specification of an often arbitrary uncertainty class. Nevertheless these two philosophies are by no means mutually exclusive. Since the identification stage of an adaptive system provides a model with some uncertainty (commonly approaching zero asymptotically) it is conjectured that a minimax processing system could aid in improving the dynamics and region of convergence of the overall adaptive system.

A terminological comment is in order here. Zadeh [33] identifies adaptivity with robustness, stating that a system is adaptive with respect to an uncertainty class if it performs acceptably well for all its members. The term robustness as referred to insensitivity to deviations in the model, has an established usage in the areas of statistics and control systems, and has been used with a minimax connotation in signal filtering problems.

The application of the minimax philosophy of game theory to filtering problems in communication can be traced back to the works of Yovits and Jackson [35] and Nilsson [36], dealing with signal estimation and matched filtering. However, the statistical works of Huber in estimation [30] and hypothesis testing [31] can be considered as the starting point of the area of minimax robustness, successfully applied to a long sequel of problems in statistical communication theory (surveyed in [34],[38],[39]). Basically, minimax filtering has been applied to three classes of linear filtering problems:

i) Wiener filtering: [26]-[28],[35],[40],[41],[43],[46],[47]

ii) Matched filtering: [3],[5],[36],[37],[43]-[45],[48]

iii) Kalman filtering: [10]-[14],[19],[49]

In this thesis we present in the first part several new general results (not specific to any particular filtering situation) and in the second part their application to various problems in linear and non-linear filtering.

The cases in minimax robust filtering for which there exists an amenable analytical solution are those for which *saddle points* exist. A filter H and an operating point P are said to form a saddle point if, fixing P, any other filter different from H has worse performance, i.e. H is the optimal filter for P, and if, fixing H, any other operating point different from P gives better performance, i.e. H has its worst performance when P is present. If there exists such a filter H, then it is the sought-after minimax robust filter, because its worst-case performance is attained at P and any other filter has worse behavior at P. Further, suppose that we use optimal filters for every operating point in the uncertainty class, then P is the element whose filter achieves the worst optimal performance, and hence will be referred to as the *least favorable* operating point. Note that the saddle point property is not necessary for the robust filter to exist; however, if it holds, the robust filter has the convenient feature of being the optimal filter for one of the operating points (the least favorable) and of performing better at any other point. In this case the minimax robust filtering problem is reduced to the search for a least favorable, since we usually suppose that the derivation of the optimal filter for a particular operating point is given by the classical theory. Frequently, there exist analytical expressions for the optimal performance achievable at every operating point; therefore the search for the least favorable turns out to be a minimization problem. At this point, the next step is to ask for conditions under which a saddle point exists. This is the question addressed by the minimax theorems of game theory (see e.g. [32]); a common assumption in these is the concavity-convexity of the performance function in the set of filters and in the set of operating points, respectively. However it is often the case that the filtering situations that we face do not *satisfy this assumption*. In the robust filtering theorem of Chapter 2, sufficient conditions are provided under which the robust filter is the optimal filter for the least favorable operating point. Roughly, these conditions are the convexity of the uncertainty set, a continuity-type condition on the performance function, and the convexity of this function in the set of operating points. However this by itself does not assure the existence of a saddle point since the uncertainty class may not have a least favorable element; in this respect the robust filtering theorem cannot be considered a minimax theorem since it leaves open for every particular uncertainty set the existence of least favorable operating points (actually, the condition of requiring uncertainty classes with least favorables is relaxed, and more general classes are allowed). The next general result presented in Chapter 2 deals with the frequent situation in which there are several independent uncertainties; under some restrictions on the performance function necessary and sufficient conditions are given for a filter and operating point pair to be a saddle point. These conditions, which appear in recursive form (and do not require convexity of the uncertainty classes), are used successfully in solving particular problems in the following chapters, and turn out to be very useful in testing candidates for saddle points. Another idea that is investigated in Chapter 2 is the soft minimax philosophy that can be employed when the modeling uncertainty is diminished by the presence of a nominal, most likely, element.

In Chapter 3, the general robust matched filtering problem is solved using the aforementioned results; in particular the uncertainties separation theorem results in a set of simpler and more general conditions than those previously published. Also, least favorable signals and noise for several useful uncertainty classes are obtained in the discrete-time case, and the optimal design of signals to combat their uncertainty in reception is investigated as well.

The problem of designing linear observers and regulators for stochastic systems in both continuous and discrete-time with uncertain second order statistics is addressed in Chapter 4. Once again, in this case the application of the general results allows the conjoint resolution of the estimation and control problems. The soft minimax solution of the unknown first-order statistics case is shown to be equivalent to earlier works in the tracking-evasion games. Particular uncertainty classes of covariance matrices are studied and some technical errors in a treatment of minimax Kalman filtering which appeared in the literature are pointed out.

In Chapter 5, several further applications are considered. Some general results for Wiener filtering are provided and a critique of the usually employed penalty function is presented. Next the robust filtering theorem and the uncertainties separation theorem are applied to two problems in random-signal detection, namely the quadratic receiver and the output-energy filter. Finally, a non-filtering application is demonstrated by providing a minimax result for hypothesis testing and an alternative proof of the classical minimax robust location estimation theorem.

2. GENERAL RESULTS

2.1 Formulation

In an unconstrained sense, given some external situation (operating point), and a processing device (filter), a quantitative measure of performance is assigned. The classical filtering problem is to find a filter that maximizes that figure of merit for a particular operating point. As discussed in the previous chapter, when the operating point is not exactly known, one of the alternatives is to search for a (maximin) filter with an optimum lower bound of performance. The goal of this chapter is to provide some general tools that can be applied in a wide variety of these filtering situations with uncertain operating points. Denote by $\mathcal{H}$ the space of filters and by $\mathcal{Q}$ the space of operating points. The payoff function M is a real functional

$$M(\cdot,\cdot): \ \mathcal{H} \times \mathcal{Q} \to \mathbb{R}$$

The triple $(\mathcal{H},\mathcal{Q},M)$ defines the type of filtering situation. Suppose $H \subset \mathcal{H}$ and $Q \subset \mathcal{Q}$ are the sets of allowable filters and operating points respectively. According to the standard terminology the triple (H,Q,M) will be referred to as a game. The following definitions will be used:

$$M^*(q) = \sup_{h \in H} M(h,q); \tag{2.1}$$

$h^*(q)$ is the <u>optimal filter</u> for $q \in Q$ if

$$M(h^*(q),q) = M^*(q); \tag{2.2}$$

q_L is a <u>least favorable</u> operating point for (H,Q,M) if

$$q_L = \arg \min_{q \in Q} M^*(q) \tag{2.3}$$

(Note that throughout this thesis we use the notation $x_0 = \arg \max/\min_{x \in A} f(x)$ if $f(x_0) = \max/\min_{x \in A} f(x)$, without implying neither existence nor uniqueness.)

$(h_L, q_L) \in H \times Q$ is a regular pair for (H,Q,M) if for every $q \in Q$ such that $x_\alpha = (1-\alpha)q_L + \alpha q \in Q$ for all $0 \leq \alpha \leq 1$,

$$\lim_{\alpha \downarrow 0} \frac{M^*(x_\alpha) - M(h_L, x_\alpha)}{\alpha} = 0 \tag{2.4}$$

$(h_L, q_L) \in H \times Q$ is a saddle point solution to the game (H,Q,M) if for every $h \in H$ and $q \in Q$

$$M(h, q_L) \leq M(h_L, q_L) \leq M(h_L, q) \tag{2.5}$$

h_R is a maximin robust filter for the game (H,Q,M) if

$$h_R = \arg \max_{h \in H} \inf_{q \in Q} M(h,q) \tag{2.6}$$

2.2 Robust filtering theorem

In this section we give some general results pertaining to the solution of the problem described by (2.6).

Lemma 2.1

Let $f: [0,1] \to \mathbb{R}$ be a convex function. Then $f(0) \leq f(\alpha)$ $\forall \alpha \in [0,1]$ if and only if $\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [f(\alpha) - f(0)]$ exists, is finite and nonnegative.

Proof

This is an extension of the proof given in [1] for an open interval.

First we show that $\frac{1}{x}$ [f(x) - f(0)] is monotone non-decreasing in $x \in (0,1]$. Suppose $0 < x \leq x'$, because $f(\cdot)$ is convex

$$f(x) \leq f(x') \cdot \frac{x}{x'} + f(0) \frac{x' - x}{x'}$$

or equivalently

$$\frac{1}{x} [f(x) - f(0)] \leq \frac{1}{x'} [f(x') - f(0)]$$

Now we can prove the only if part:

Suppose $f(0) \leq f(\alpha)$ $\forall \alpha \in [0,1]$, then $\forall \alpha \in (0,1]$, $\frac{1}{\alpha} [f(\alpha) - f(0)] \geq 0$, but since this function is monotone non-decreasing $\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [f(\alpha) - f(0)]$ exists, is finite and nonnegative.

For the reverse implication, assume that $\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [f(\alpha) - f(0)]$ exists, is finite and nonnegative, then for every $x \in (0,1]$, $\exists$ $z \leq x$ such that $\frac{1}{z} [f(z) - f(0)] \geq 0$, but since this function is monotone non-decreasing in z, $\frac{1}{x} [f(x) - f(0)] \geq 0$, hence $f(x) \geq f(0)$.

Lemma 2.2

Suppose Q is a convex set, and for every $h \in H$, $M(h,\cdot)$ is convex in Q. With respect to the game (H,Q,M), a regular pair (h_L,q_L) is a saddle point solution if and only if q_L is least favorable.

Proof

The left inequality in the definition of saddle point (2.5) is satisfied for every regular pair, because particularizing the regularity condition (2.4) for $q = q_L$, it follows that

$$M^*(q_L) = M(h_L,q_L) \tag{2.7}$$

According to Lemma 2.1, the right inequality is satisfied if and only if

$$A(q) = \lim_{\alpha \downarrow 0} \frac{M(h_L,x_\alpha) - M(h_L,q_L)}{\alpha} \geq 0 \tag{2.8}$$

for every $q \in Q$.

Now, note that $M^*(\cdot)$ is convex in Q since with $q_0 = (1-\alpha)q_1 + \alpha q_2$, for all $0 \leq \alpha \leq 1$, and $q_1,q_2 \in Q$, we have

$$\begin{aligned}\sup_{h \in H} M(h,q_0) &\leq \sup_{h \in H} \{(1-\alpha)M(h,q_1) + \alpha M(h,q_2)\} \\ &= (1-\alpha) \sup_{h \in H} M(h,q_1) + \alpha \sup_{h \in H} M(h,q_2)\end{aligned} \tag{2.9}$$

where the inequality follows from the assumed convexity of M in Q for every $h \in H$.

Again by Lemma 2.1, since $M^*(\cdot)$ is convex in Q, q_L is a least favorable operating point if and only if

$$B(q) = \lim_{\alpha \downarrow 0} \frac{M^*(x_\alpha) - M^*(q_L)}{\alpha} \geq 0 \tag{2.10}$$

for every $q \in Q$.

Considering that

$$M^*(x_\alpha) - M^*(q_L) = M^*(x_\alpha) - M(h_L,x_\alpha) + M(h_L,x_\alpha) - M(h_L,q_L) \tag{2.11}$$

and that (q_L,h_L) is a regular pair, taking $\lim_{\alpha \downarrow 0} \frac{1}{\alpha}[\quad]$ of both sides of 2.11, we get that for any $q \in Q$

$$A(q) = B(q) , \tag{2.12}$$

and therefore the lemma is proved.

Theorem 2.1

Suppose that (K,P,M) is a game with the following properties:

i) There exists a convex set Q, $P \subset Q \subset \mathcal{Q}$ such that

$$\inf_{q \in P} \sup_{h \in K} M(h,q) = \inf_{q \in Q} \sup_{h \in K} M(h,q) \tag{2.13}$$

ii) There exists a set H, such that $H \subset K \subset \mathcal{K}$ and for every $q \in Q$

$$\sup_{h \in K} M(h,q) = \sup_{h \in H} M(h,q)$$

iii) $M(h,\cdot)$ is convex in Q for all $h \in H$.

Then, if a least favorable for (H,Q,M) q_L, and its optimal filter h_L form a regular pair for (H,Q,M), h_L is a maximin robust filter for the game (K,P,M).

Proof

According to Lemma 2.2, if a least favorable for (H,Q,M), q_L, and its optimal filter, h_L, form a regular pair, (h_L,q_L) is a saddle point solution to the game (H,Q,M). Therefore, we have

$$\begin{aligned} M(h_L,q_L) &= \inf_{q \in Q} M(h_L,q) \\ &\leq \inf_{q \in P} M(h_L,q) \\ &\leq \sup_{h \in K} \inf_{q \in P} M(h,q) \end{aligned} \tag{2.14}$$

where the first inequality follows from $P \subset Q$ and the second from $h_L \in H \subset K$. Furthermore,

$$
\begin{aligned}
M(h_L,q_L) &= \sup_{h \in H} M(h,q_L) \\
&\geq \inf_{q \in Q} \sup_{h \in H} M(h,q) \\
&= \inf_{q \in P} \sup_{h \in K} M(h,q) \qquad (2.15)
\end{aligned}
$$

where the inequality follows from $q_L \in Q$ and the last equality from assumptions i) and ii). But for an arbitrary game we have (e.g. [2])

$$
\sup_{h \in K} \inf_{q \in P} M(h,q) \leq \inf_{q \in P} \sup_{h \in K} M(h,q) \qquad (2.16)
$$

Therefore the previous inequalities in (2.14) and (2.15) are transformed into equalities. In particular,

$$
\sup_{h \in K} \inf_{q \in P} M(h,q) = \inf_{q \in P} M(h_L,q) \qquad (2.17)
$$

and therefore h_L is a maximin robust filter for (K,P,M), as we wanted to show.

Remarks

The utility of this result stems from the fact that usually explicit expressions for $M^*(\cdot)$ are available, and therefore the original maximinimization problem is reduced to a minimization one, namely the search for least favorables. This will be illustrated extensively in the next chapters with the application of the general results to specific filtering situations.

It seems convenient to underline the meaning of the introduction of a game (H,Q,M) different from the original (K,P,M) in the theorem. The key point is the existence of a game with saddle point, in order that Lemma 2.2 can be used. On one hand, the introduction of the set Q allows the solution of problems in which the original uncertainty set either is not convex or has no least favorable. Note that in the case in which $\inf_{q \in Q} \sup_{h \in K} M(h,q)$ is attained, the existence of a robust filter is assured. On the other hand, the introduction of the set H allows the solution of the robust filtering problem for general sets of filters (class theorems), while dealing with easier, restricted sets.

However, the most relevant remark is that important robust filtering situations (as will be seen in the next chapters) cannot be solved by using the well-known game-theoretic minimax theorems, because their respective payoff functions, although convex in the uncertainties, are not concave in the filter sets.

2.3 Uncertainties separation theorem

Let (H,Q,M) be a game in which Q is the cartesian product of independent uncertainty classes

$$Q = Q_1 \times Q_2 \times \dots \times Q_k$$

and the payoff function can be put as

$$M(h,(q_1,\dots,q_k)) = F(f_1(h,q_1),\dots,f_k(h,q_k)) \tag{2.18}$$

with $F: \mathbb{R}^k \to \mathbb{R}$ non-decreasing in each one of its arguments.

Define, if they exist, for $i = 1,\dots,k$

$$q_i^*(h) = \arg \min_{x \in Q_i} f_i(h,x) \tag{2.19}$$

Theorem 2.2

(h_L,q_L) is a saddle point of (H,Q,M) if and only if h_L is solution of the equation

$$h_L = \arg \max_{h \in H} M(h, (q_1^*(h_L), \ldots, q_k^*(h_L))) \tag{2.20}$$

and

$$q_L = (q_1^*(h_L), \ldots, q_k^*(h_L))$$

<u>Proof</u>

Only if. Suppose (h_L, q_L) is a saddle point of (H,Q,M), then

$$M(h_L, q_L) = \min_{q \in Q} M(h_L, q) = \min_{q \in \hat{Q}} M(h_L, q) \tag{2.21}$$

for all $\hat{Q} \subset Q$, such that $q_L \in \hat{Q}$. In particular for $Q_i = \{q_{L1}\} \times \{q_{L2}\} \times \ldots \times Q_i \times \ldots \times \{q_{Lk}\}$ with $q_L = (q_{L1}, \ldots, q_{Lk})$. This leads to

$$M(h_L, q_L) = \min_{x \in Q_i} M(h_L, (q_{L1}, \ldots, x, q_{L(i+1)}, \ldots, q_{Lk})) \tag{2.22}$$

Equivalently

$$F(f_1(h_L, q_{L1}), \ldots, f_i(h_L, q_{Li}), \ldots, f_k(h_L, q_{Lk})) =$$

$$= \min_{x \in Q_i} F(f_1(h_L, q_{L1}), \ldots, f_i(h_L, x), \ldots, f_k(h_L, q_{Lk})) \tag{2.23}$$

and since F is non-decreasing in the i-th argument

$$q_{Li} = \operatorname{argmin}_{x \in Q_i} f_i(h_L, x) = q_i^*(h_L) \tag{2.24}$$

Moreover, if (h_L, q_L) is a saddle point,

$$h_L = \operatorname{argmax}_{h \in H} M(h, q_L) \tag{2.25}$$

Then, the last two equations result in (2.20).

If. From the definition of $q_i^*(h_L)$ and using the monotonicity of F in every argument we have that

$$M(h_L,q_L) = M(h_L,(q_1^*(h_L),\ldots,q_k^*(h_L))$$
$$= \min_{q \in Q} M(h,q) \tag{2.26}$$

Furthermore, (2.20) is the remaining condition in order for (h_L,q_L) to be a saddle point of (H,Q,M).

Corollary

Suppose the game (K,P,M) satisfies the conditions stated in Theorem 2.1, and h_L is solution to the equation (2.20), then h_L is a maximin robust filter for the game (K,P,M).

Proof

Letting $q_L = (q_1^*(h_L),\ldots,q_k^*(h_L))$, (h_L,q_L) is a saddle point of (H,Q,M). Then the same proof of Theorem 2.1 can be used.

Remarks

Although Theorem 2.2 is quite apparent (for k = 1 it is nothing else than the definition of saddle point), it represents a powerful tool in order to derive conditions for least favorability, and in order to check candidates for saddle points.

This result is especially attractive for an iterative numerical solution of the robust filtering problem, i.e., for a given filter the solution of the set of equations (2.19) can be computed and incorporated into equation (2.20) which in turn can be solved for a new filter. Furthermore, if the uncertainty subclasses are such that there exists an analytical solution (in terms of a generic h) for $q_i^*(h)$, $i = 1,\ldots,k$, then every uncertainty can be solved separately (as if all the others were constant), getting finally the robust filter from the (also, frequently analytic) solution of (2.20).

One particular case in which there exists an explicit solution for the least favorable and that can be found repeatedly in the next chapters is given by the following result.

Lemma 2.3

Suppose that the uncertainty subclass Q_i is a subset of a Hilbert space S and is some neighborhood around a nominal element described by

$$Q_i = \{x \in S,\ \|x - x_0\| \le \Delta\} \tag{2.27}$$

where the norm is that of the Hilbert space S. Suppose that $f_i(h_L,\cdot): S \to \mathbb{R}$ is a continuous linear functional, then

$$q_i^*(h_L) = x_0 - \Delta \frac{z(h_L)}{\|z(h_L)\|} \tag{2.28}$$

where $z(h_L)$ is such that for all $x \in S$

$$f_i(h_L,x) = \langle x, z(h_L)\rangle \tag{2.29}$$

Proof

Since $f_i(h_L,\cdot)$ is a continuous linear functional the existence of an element $z(h_L) \in S$ fulfilling (2.29) is assured by the Riesz representation theorem (e.g. [8]). Note that for all $x \in Q_i$ the Schwarz inequality results in

$$|\langle x-x_0, z(h_L)\rangle| \le \|x - x_0\|\,\|z(h_L)\| \le \Delta\,\|z(h_L)\| \tag{2.30}$$

Also,

$$\langle x-(x_0-\Delta\frac{z(h_L)}{\|z(h_L)\|}),z(h_L)\rangle = \langle x-x_0,z(h_L)\rangle + \Delta\|z(h_L)\| \geq 0 \qquad (2.31)$$

where the inequality follows from (2.30). Therefore we have for all $x \in Q_i$

$$f_i(h_L,x) \geq f_i(h_L, x_0-\Delta\frac{z(h_L)}{\|z(h_L)\|}) \qquad (2.32)$$

and (2.28) follows.

2.4 Soft minimax filtering

Sometimes there exists a nominal operating point q_0, such that our uncertainty on the actual operating point is lessened by the fact that points closer, in some sense, to the nominal are more likely to occur than those more distant from it. This can be incorporated into our model, modifying the payoff function in such a way that points in the uncertainty class are penalized according to some distance from the nominal ($D(q,q_0)$). Hence for soft minimax, we have the payoff function

$$M_s(h,q) = M(h,q) + D(q,q_0) \qquad (2.33)$$

Note that now, the elements closer to q_0 are stronger candidates to be the worst case. All the results derived before can be applied to the game with payoff function M_s. Directly from the definition of regularity we can see that (h_L,q_L) is a regular pair for (H,Q,M) if and only if it is for (H,Q,M_s). Besides it is not unusual that the distance function is convex in the uncertainties; if so, condition iii) in Theorem 2.1 is also equivalent for both types of minimax filtering. However the existence of a least favorable for (H,Q,M) does not guarantee its existence for (H,Q,M_s), even when D is convex. An example is sufficient in order to verify this.

Consider the convex set

$$Q = \{(x_1,x_2): \ 0 < x_1, \ x_2 < 1\} \cup \{(0,1)\} \cup \{(1,0)\} \subset \mathbb{R}^2$$

and the convex functions $f(\underline{x}) = x_1$, $g(\underline{x}) = x_2$. Obviously

$$\min_{\underline{x} \in Q} f(\underline{x}) = f((0,1))$$

$$\min_{\underline{x} \in Q} g(\underline{x}) = g((1,0))$$

But there exists no $\underline{z} \in Q$ such that

$$f(\underline{z}) + g(\underline{z}) = 0 = \inf_{\underline{x} \in Q} \{f(\underline{x}) + g(\underline{x})\}$$

3. MATCHED FILTERING

3.1 Robust Matched Filtering Theorems

The linear system that maximizes the output signal-to-noise ratio at some instant of time, when the input is a deterministic signal embedded in additive random noise is known as the matched filter for this pair of signal and noise. If the noise is a Gaussian process, then the output of this filter in the instant in which the signal-to-noise ratio is maximized provides a sufficient statistic for any likelihood ratio detection test of the input signal. Since the power of the noise at the output of the linear filter depends on the second-order statistics of the input noise, a complete specification of the signal and autocorrelation of the noise is necessary and sufficient in order to derive the corresponding matched filter. Due to modeling uncertainties or changing operating environments it is possible that the second order characterization of the noise is not completely known. Also, channel nonlinearities tend to distort the signal in an unpredictable (or difficult to ascertain) way. In these cases it is interesting to design a robust matched filter, i.e. a filter that gives the optimum - in some sense - behavior within the uncertainty region. Poor [3] showed that under some mild restrictions the design of the robust matched filter in the maximin sense - the most widely used in robust decision problems - is equivalent to finding the least favorable possible pair of signal and noise. In this chapter we derive the central results of robust matched filtering by using the theorems proved in Chapter 2, and present some analytical solutions for particular uncertainty classes in the increasingly important case of discrete time processing [4].

A general formulation of the matched filter design problem that allows the description of the input pair of signal and noise in various ways has been given in [3]. Let a signal quantity (in the time or frequency domain) be $s \in \mathcal{H}$, a noise quantity (e.g., covariance matrix or autocorrelation function) be $n \in \bar{\mathcal{H}}$, and a filter quantity (e.g. impulse response or transfer function) be $h \in \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space with inner product $\langle \cdot,\cdot \rangle$ and $\bar{\mathcal{H}}$ is a space of bounded, linear, (self-adjoint) positive operators mapping $\mathcal{H}$ to itself. The real valued functional defined by

$$\rho(h;s,n) = |\langle h,s \rangle|^2 / \langle h,nh \rangle \tag{3.1}$$

represents the output signal-to-noise ratio of the filter at some time instant, for the usual descriptions of signal and noise, either in continuous or discrete time. Note that in order for this definition to make sense, the quantity n should represent a *second order* characterization of the noise. Therefore, the filtering situation defining triple is $(\mathcal{H},\mathcal{H} \times \bar{\mathcal{H}},\rho)$. By direct application of the Schwarz inequality the optimal (matched) filter for (s,n) is such that $n \cdot h^*(s,n) = s$. Parallel to the notation in (2.1), we write

$$\rho^*(s,n) = \sup_{h \in \mathcal{H}} \rho(h;s,n) = \langle s, h^*(s,n) \rangle \tag{3.2}$$

Now suppose that the signal and noise pair (s,n) belongs to an uncertainty class P. (A particular case is when the uncertainties in signal and noise are independent in which case $P = \mathcal{S} \times \mathcal{N}$). Thus we have the game $(\mathcal{H},P,\rho)$. A pair of signal and noise $(s_L,n_L) \in P$ is said to be least favorable for matched filtering if

$$(s_L,n_L) = \arg \min_{(s,n) \in P} \rho^*(s,n) \tag{3.3}$$

Furthermore, we will assume, in order to simplify the derivations that for every $(s,n) \in P$ the optimal filter $h^*(s,n) \in \mathcal{H}$ exists. (Note that if n is positive, it is not necessarily invertible; however, if $h^*(s,n)$ exists it is uniquely defined.) The central result in maximin robust matched filtering is the following.

<u>Theorem 3.1</u>

Suppose that

1. The uncertainty set P is convex.
2. $h^*(s,n)$ is defined for all $(s,n) \in P$.
3. $\langle s, h^*(s_L, (1-\alpha)n_L + \alpha n)\rangle$ is right continuous at $\alpha = 0$ for all $(s,n) \in P$.

Then, if (s_L,n_L) is a least favorable pair, its optimal filter is a robust matched filter for $(\mathcal{H},P,\rho)$.

<u>Proof</u>

According to Theorem 2.1, since $\rho(h;s,n)$ is convex in P for every $h \in \mathcal{H}$ [3] (note, incidentally, that for every $(s,n) \in P$, the SNR is not concave in $\mathcal{H}$), all we have to prove is that $(h_L,(s_L,n_L))$ is a regular pair, where $n_L h_L = s_L$.

Let $(\sigma,\nu) = (1-\alpha)(s_L,n_L) + \alpha(s,n)$, with $(s,n) \in P$. (By the assumed convexity of P, $(\sigma,\nu) \in P$.) Noting that the operator $h^*(\cdot,\nu)$ is linear we get the following equation for later use

$$
\begin{aligned}
&(1-\alpha)[\langle s_L, h^*(s_L,\nu)\rangle - \langle s_L,h_L\rangle] = \\
&= \langle s_L, h^*(s_L - \nu h_L, \nu)\rangle - \alpha\langle s_L, h^*(s_L,\nu) - h_L\rangle = \\
&= \alpha[\langle s_L, h^*(s_L - nh_L,\nu) - h^*(s_L,\nu) + h_L\rangle] = \\
&= \alpha[\langle s_L, h_L\rangle - \langle s_L, h^*(nh_L,\nu\rangle] \qquad (3.4)
\end{aligned}
$$

Now we compute

$$\rho^*(\sigma,\nu) - \rho(h_L;\sigma,\nu) = \frac{1}{\langle h_L,\nu h_L\rangle}[\langle\sigma,h^*(\sigma,\nu)\rangle\langle h_L,\nu h_L\rangle - |\langle\sigma,h_L\rangle|^2] \tag{3.5}$$

The numerator can be expanded in the expression

$$\begin{aligned}
&[(1-\alpha)^2\langle s_L,h^*(s_L,\nu)\rangle + \alpha^2\langle s,h^*(s,\nu)\rangle + \\
&\alpha(1-\alpha)(\langle s,h^*(s_L,\nu)\rangle + \langle s_L,h^*(s,\nu)\rangle)]\,[\langle h_L,\nu h_L\rangle] - \\
&[(1-\alpha)^2(\langle s_L,h_L\rangle)^2 + \alpha^2|\langle s,h_L\rangle|^2 + 2\alpha(1-\alpha)\langle s_L,h_L\rangle \operatorname{Re}\langle s,h_L\rangle] \\
= {}&(1-\alpha)^2\langle s_L,h_L\rangle[\langle s_L,h^*(s_L,\nu)\rangle - \langle s_L,h_L\rangle] + \\
&2\alpha\langle s_L,h_L\rangle[\operatorname{Re}\langle s,h^*(s_L,\nu)\rangle - \operatorname{Re}\langle s,h_L\rangle] + \\
&\alpha\langle s_L,h^*(s_L,\nu)\rangle(\langle h_L,nh_L\rangle - \langle h_L,n_Lh_L\rangle) + o(\alpha)
\end{aligned} \tag{3.6}$$

Where we have used that $h^*(\cdot,\nu)$ is self-adjoint. The first term of the right hand of (3.6) is given by (3.4); therefore, the numerator of (3.5) is

$$\begin{aligned}
&(1-\alpha)\alpha\,[\langle s_L,h_L\rangle - \langle s_L,h^*(nh_L,\nu)\rangle]\langle s_L,h_L\rangle + \\
&2\alpha\langle s_L,h_L\rangle[\operatorname{Re}\langle s,h^*(s_L,\nu)\rangle - \operatorname{Re}\langle s,h_L\rangle] + \\
&\alpha\langle s_L,h^*(s_L,\nu)\rangle(\langle h_L,nh_L\rangle - \langle s_L,h_L\rangle) + o(\alpha) \\
= {}&\alpha\,[\langle s_L,h_L\rangle\langle s_L,h_L - h^*(s_L,\nu)\rangle + \langle s_L,h^*(s_L,\nu)\rangle\langle h_L,nh_L\rangle - \\
&\langle s_L,h_L\rangle\langle h^*(s_L,\nu),nh_L\rangle] + o(\alpha)
\end{aligned} \tag{3.7}$$

Using assumption 3, the term in brackets goes to zero when $\alpha \downarrow 0$, therefore (3.5) is $o(\alpha)$ and consequently (s_L,n_L) is a regular pair since (s,n) is arbitrary.

Remarks

For the sake of clarity the theorem is not stated in its most general form. It is enough that the uncertainty class P has property i) of Theorem 2.1. With respect to the continuity condition required by the theorem, we have the following

Lemma 3.1

If n_L is invertible then

$$\lim_{\alpha \downarrow 0} h^*(s_L,\nu) = h^*(s_L,n_L)$$

for all $n \in \bar{\mathcal{K}}$ such that $(s_L,n) \in P$, where $\nu = (1-\alpha)n_L + \alpha n$.

Proof

$$\|\nu x\| = \|((1-\alpha)n_L + \alpha n)x\| \geq (1-\alpha)\|n_L x\| - \alpha\|nx\|$$

$$\geq (1-\alpha)\|n_L x\| - \alpha\|n\|\|x\| \tag{3.8.1}$$

where the last inequality follows because n is bounded. If n_L is invertible (one to one and onto) then (Theorem 21.3 [21]) there exists $\varepsilon > 0$ such that for every vector x

$$\|n_L x\| \geq \varepsilon\|x\|$$

Now fix t such that $0 < t < \varepsilon/(\varepsilon + \|n\|)$. Then according to (3.8.1) for every $\alpha \in [0,t]$ and every x,

$$\|\nu x\| \geq \delta\|x\| \tag{3.8.2}$$

with $\delta = (1-t)\varepsilon - t\|n\| > 0$.

Since ν is self-adjoint positive

$$\bar{\mathcal{H}} = \text{Null }(\nu)^{\perp} = \overline{\text{Range }(\nu)}$$

therefore, the range of ν is dense. This fact and (3.8.2) imply that ν is invertible (Theorem 21.3 [21]). If we let $x = \nu^{-1}y$ in (3.82) for every vector y we have

$$\|y\| = \|\nu\nu^{-1}y\| \geq \delta\|\nu^{-1}y\|$$

and therefore for every $\alpha \in [0,t]$

$$\|\nu^{-1}\| = \sup_{y} \frac{\|\nu^{-1}y\|}{\|y\|} \leq \frac{1}{\delta} \tag{3.8.3}$$

The existence of ν^{-1} (for $\alpha \in [0,t]$) implies that the following is true

$$h^*(s_L,\nu) - h^*(s_L,n_L) = \nu^{-1}(n_L - \nu)h^*(s_L,n_L)$$

Taking norms

$$\begin{aligned} \|h^*(s_L,\nu) - h^*(s_L,n_L)\| &\leq \|\nu^{-1}\| \, \|n_L - \nu\| \, \|h^*(s_L,n_L)\| \\ &\leq \alpha \cdot \frac{1}{\delta} \|n_L - n\| \, \|h^*(s_L,n_L)\| \end{aligned} \tag{3.8.4}$$

and the lemma is proved.

Note that, by the Banach inverse theorem [9], since ν is linear and continuous its inverse (it is exists) is bounded. However this would not be enough for the ¬revious proof, since it requires that ν^{-1} is uniformly bounded in a neighborhood of $\alpha = 0$.

The conclusion is that the invertibility of n_L is sufficient for condition 3 to be true (via Lemma 3.1 and Schwarz inequality).

Now we are going to direct our attention to the important special case in which the uncertainties in signal and noise are independent of each other and therefore $P = \mathcal{S} \times \mathcal{N}$.

Theorem 3.2

$(h_L,(s_L,n_L))$ is a saddle point of $(\mathcal{H},\mathcal{S}\times\mathcal{N},\rho)$ if and only if

1. $n_L h_L = s_L$
2. $|\langle s,h_L\rangle| \geq |\langle s_L,h_L\rangle| \quad \forall s \in \mathcal{S}$ (3.9)
3. $\langle h_L,(n_L - n)h_L\rangle \geq 0 \quad \forall n \in \mathcal{N}$ (3.10)

Proof

The payoff function for matched filtering (SNR), (3.1), can be expressed as

$$\rho(h;s,n) = F(|\langle s,h\rangle|,(\langle h,nh\rangle)^{-1}) \tag{3.11}$$

with $F\colon \mathbb{R}^2 \to \mathbb{R}$, $F(x,y) = x^2y$. Consequently the uncertainties separation theorem can be applied and equations (2.20) and (2.19) result in conditions 1, and 2-3 straightforwardly.

Remark

In [3] a similar result (for convex classes) was derived with conditions 2 and 3 replaced by

$$2\,\mathrm{Re}\,\{\langle s,h_L\rangle\} - \langle s_L,h_L\rangle - \langle h_L,nh_L\rangle \geq 0 \tag{3.12}$$

for every $(s,n) \in \mathcal{S}\times\mathcal{N}$. In [5] it was shown that this is equivalent to condition 3 and

$$\mathrm{Re}\,\{\langle s,h_L\rangle\} \geq \langle s_L,h_L\rangle \quad \forall s \in \mathcal{S} \tag{3.13}$$

It is interesting to note that if S is convex (Theorem 3.2 does not put any restriction on S and N) (3.9) and (3.13) are equivalent. First, it is obvious that (3.13) implies (3.9). Now suppose there exists $s_1 \in S$ such that $\mathrm{Re}\{\langle s_1,h_L\rangle\} < \langle s_L,h_L\rangle$ and consider $\sigma_\alpha = (1-\alpha)s_L + \alpha s_1$.

$$\begin{aligned}|\langle \sigma_\alpha,h_L\rangle|^2 &= [(1-\alpha)\langle s_L,h_L\rangle + \alpha\ \mathrm{Re}\ \langle s_1,h_L\rangle]^2 + \\ &\quad [\alpha\ \mathrm{Im}\ \langle s_1,h_L\rangle]^2 \\ &= (\langle s_L,h_L\rangle)^2 + \alpha^2(\mathrm{Re}\ \langle s_1,h_L\rangle - \langle s_L,h_L\rangle)^2 + \\ &\quad 2\alpha\ \langle s_L,h_L\rangle\ (\mathrm{Re}\ \langle s_1,h_L\rangle - \langle s_L,h_L\rangle) + \\ &\quad \alpha^2(\mathrm{Im}\ \langle s_1,h_L\rangle)^2 \end{aligned} \tag{3.14}$$

Denoting $\Delta = \langle s_L,h_L\rangle - \mathrm{Re}\ \langle s_1,h_L\rangle > 0$, for every α such that

$$0 < \alpha \le 2\Delta\ \langle s_L,h_L\rangle/(\Delta^2 + (\mathrm{Im}\ \langle s_1,h_L\rangle)^2) \tag{3.15}$$

we have that $|\langle \sigma_\alpha,h_L\rangle|^2 < |\langle s_L,h_L\rangle|^2$. Therefore, since S is convex, $\sigma_\alpha \in S$ and (3.9) does not hold.

3.2 Application to Discrete Time

One of the most important instances in which this problem appears is in Direct Sequence Spread Spectrum systems. In this case, the optimum linear processor is an analog filter matched to the entire codeword waveform. However, because of technological or versatility considerations it is often preferred to use a discrete-time matched filter whose input can proceed from a direct sampling of the received codeword or from the sampling of the output of an analog filter matched to the chip waveform.

In the above formulation, let $\mathcal{H} = \mathbb{R}^k$, $s = [s_0,\ldots,s_{k-1}]^T$, $h = [h_0,\ldots,h_{k-1}]^T$, where $h_i = \tilde{h}_{k-i-1}$, and s_i, $\tilde{h}_i$ are the values of signal and filter impulse response respectively at the i-th sample. The inner product is defined as the usual scalar product: $\langle a,b \rangle = a^T b$; and the noise descriptor is $n \in \bar{\mathcal{K}} \subset \mathbb{R}^{k\times k}$, a positive-definite symmetric matrix, representing the covariance matrix of the zero-mean input noise. It is easy to see that with these definitions, (3.1) gives the power of the filter output due to the signal, divided by the variance of the filter output due to the noise, at the k-1-th sample.

Note that in this (finite dimensional) case the noise operator (assumed to be a positive definite matrix) is always invertible (thus $h^*(s,n)$ exists for every pair) and through Lemma 3.1 the continuity condition of Theorem 3.1 is always satisfied.

In the sequel, we will deal with simple and general uncertainty classes in order to illustrate the search for least favorable pairs.

3.2.1 Signal Uncertainty

We consider here two classes of uncertainties described by a bound on the ℓ_2 and ℓ_∞ norm of the deviation from a given nominal s_0.

[ℓ_2 uncertainty]. Let n_0 be the covariance matrix of the input noise, and S_1 the class of allowable signals defined by:

$$S_1 = \{s \in \mathbb{R}^k,\ \|s - s_0\| < \Delta\} \tag{3.16}$$

The least favorable signal s_L depends on n_0 and is given by the following:

Proposition 1

$$s_L = \underset{s \in S_1}{\operatorname{argmin}} \langle s, n_0^{-1} s \rangle = s_0 - \sigma^2 h_L \tag{3.17}$$

with

$$\sigma^2 \| h_L \| = \Delta \tag{3.18}$$

and $h_L = n_0^{-1} s_L$.

Proof

Since the covariance matrix of the noise is known and is positive definite, according to Theorem 3.2, s_L is least favorable if and only if

$$\langle s_L, h_L \rangle = \min_{s \in S_1} \langle s, h_L \rangle \; ; \tag{3.19}$$

but by (3.17) and (3.18) we have for $s \in S_1$

$$\langle s - s_L, h_L \rangle = \langle s - s_0, h_L \rangle + \Delta \| h_L \| \; , \tag{3.20}$$

where the right side is nonnegative by the Schwarz inequality:

$$|\langle s - s_0, h_L \rangle| \leq \| h_L \| \, \| s - s_0 \| \leq \Delta \| h_L \| \; . \tag{3.21}$$

Since $s_L \in S_1$, s_L is the least favorable signal.

We can get an alternative expression for the robust filter, with (3.17) and (3.18):

$$h_L = (n_0 + \sigma^2 I)^{-1} s_0 \tag{3.22}$$

Equation (3.22) shows that the robust filter is the filter matched to the nominal signal and the input noise with an added component of white noise of variance σ^2. Note that in general, σ is computed recursively from (3.18) and (3.22). Further simplification of the result can be obtained in particular cases such as the following:

Proposition 2

The - nominal - filter matched to (s_0,n_0) is robust for deviations from s_0 defined by the class S_1 (3.16) if and only if s_0 is an eigenvector of n_0.

Proof

First note from (3.1) that the performance of filter h is not affected by scaling the impulse response by a constant, so the nominal filter is robust if and only if there exists $k \in \mathbb{R}_0$, such that

$$(n_0 + \sigma^2 I)^{-1} s_0 = k n_0^{-1} s_0 \tag{3.23}$$

but this is equivalent to

$$[(1-k)/\sigma^2]\, s_0 = n_0^{-1} s_0 \tag{3.24}$$

and since n_0^{-1} is invertible, (3.24) holds if and only if s_0 is an eigenvector of n_0.

[ℓ_∞ uncertainty]. Let S_2, the class of allowable signals, be defined by

$$S_2 = \{s \in \mathbb{R}^k,\ \max|s_i - s_{0i}| \leq \Delta,\ i = 0,\dots,k-1\} \tag{3.25}$$

In order to get the least favorable signal in this class, according to (3.3) we have to find out the solution to the minimization problem:

$$\underset{s \in J_0 \times \dots \times J_{k-1}}{\operatorname{argmin}} \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} (n_0^{-1})_{ij}\, s_i s_j \tag{3.26}$$

with $J_i = [s_{0i} - \Delta,\ s_{0i} + \Delta]$. In some special cases an analytic result is achievable:

Proposition 3

If the samples of the noise are uncorrelated the least-favorable signal s_L in S_2 is given by:

$$s_{Li} = \begin{cases} 0 & -\Delta \leq s_{0i} \leq \Delta \\ s_{0i} - \Delta & \Delta < s_{0i} \\ s_{0i} + \Delta & s_{0i} < -\Delta \end{cases} \tag{3.27}$$

Proof

If the noise samples are uncorrelated we have:

$$n_0^{-1} = \mathrm{diag}(\ell_1, \ldots, \ell_k)$$

with $\ell_i > 0$. For all $i = 0, \ldots, k-1$, it is easy to see that:

$$s_i s_{Li} \geq (s_{Li})^2 \tag{3.28}$$

for any $s \in S_2$. Since $h_{Li} = \ell_i s_{Li}$, this implies:

$$\langle s_L, h_L \rangle = \min_{s \in S_2} \langle s, h_L \rangle \tag{3.29}$$

Therefore, by Theorem 3.2, s_L is the least favorable signal.

Proposition 4

If there exists an element $s_L \in S_2$ such that

$$s_{Li} = \begin{cases} s_{0i} - \Delta & \text{if } h_{Li} > 0 \\ s_{0i} + \Delta & \text{if } h_{Li} < 0 \end{cases} \tag{3.30}$$

with $h_L = n_0^{-1} s_L$, then s_L is the least favorable signal.

Proof

The expressions (3.25) and (3.30) imply that for any $s \in S_2$ and $i = 0,\ldots,k-1$,

$$h_{Li} s_i \geq h_{Li} s_{Li} \tag{3.31}$$

Since this is sufficient in order to have (3.29), s_L is the least favorable signal in S_2.

Note that this proposition suffers from the common inconvenience of the results derived from the uncertainties separation theorem, namely s_L depends on h_L, and therefore the solution must be reached recursively except for special cases. However, we can assure the existence of a solution of the type of (3.30) and its direct computation under the condition of the following:

Proposition 5

If the maximum deviation from the nominal signal in each sample is bounded by:

$$\Delta < \min_i |h_{0i}| / \max_j \sum_m |(n_0^{-1})_{jm}| \tag{3.32}$$

the least favorable signal s_L in S_2 is given by

$$s_{Li} = \begin{cases} s_{0i} - \Delta & \text{if } h_{0i} > 0 \\ s_{0i} + \Delta & \text{if } h_{0i} < 0 \end{cases} \tag{3.33}$$

where $h_0 = n_0^{-1} s_0$ is the nominal matched filter.

Proof

With $h_L = n_0^{-1} s_L$ and $u_L = s_L - s_0$, for $i = 0,\ldots,k-1$

$$|h_{Li}-h_{0i}| = |(n_0^{-1}u_L)_i| \tag{3.34}$$

since $s_L \in S_2$, the absolute value of the components of u_L is bounded by Δ. Thus

$$|(n_0^{-1}u_L)_i| \leq \Delta \max_j \sum_m |(n_0^{-1})_{jm}| \tag{3.35}$$

and by (3.32)

$$|h_{Li} - h_{0i}| < |h_{0i}| \tag{3.36}$$

which is sufficient in order that

$$\mathrm{sgn}(h_{Li}) = \mathrm{sgn}(h_{0i}) \tag{3.37}$$

Therefore, by Proposition 4, s_L given by (3.33) is the least favorable signal in S_2.

3.2.2 Noise Uncertainty

Here we suppose that the nominal signal s_0 is truly present at the input, but the actual covariance matrix n is allowed to differ from the nominal n_0. The first is a general result useful for different classes of uncertainties.

Lemma 3.2

n_L is the least favorable noise for every nominal signal $s_0 \in \mathbb{R}^k$ if and only if n_L is a maximal element of N (the uncertainty class).*

Proof

Since there is no uncertainty in the signal we only need to prove

---

*It is supposed that N does not depend on s_0.

conditions 1 and 3 of Theorem 3.2.

By definition, n_L is a maximal element of N, if and only if

$$n_L \geq n \tag{3.38}$$

for all $n \in N$. But this is equivalent to (3.10) holding for all $s_0 \in \mathbb{R}^k$ since for every $h_L \in \mathbb{R}^k$ there exists $s_0 = n_L h_L$.

As an application of this previous Lemma, we have a result on the least favorability of white noise:

Proposition 6

Suppose $n_L = cI \in N$, this is the least favorable covariance in N if and only if $\|n\|_2{}^* \leq c$ for all $n \in N$.

Proof

According to Lemma 3.2, cI is the least favorable covariance in N if and only if it is the maximal element of the class, i.e. for all $n \in N$, $x \in \mathbb{R}^k$, $x \neq 0$:

$$x^T(n_0 - n)x = c\|x\|^2 - x^T n x \geq 0 \tag{3.39}$$

$$(x^T n x)/\|x\|^2 \leq c \tag{3.40}$$

Equivalently, by Rayleigh's principle [6],

$$\rho(n) \leq c \tag{3.41}$$

where $\rho(n)$ denotes the spectral radius of n (maximum absolute value of its eigenvalues), but $\|n\|_2 = \rho(n)$.

$^*\|n\|_2$ = Operator norm of n.

Next, in analogy with the ℓ_2 signal uncertainty, we deal with a specific deviation class defined by

$$N_1 = \{n \in \mathbb{R}^{k \times k},\ \|n - n_0\| \leq \varepsilon,\ n > 0\} \tag{3.42}$$

where ε is a positive constant, n_0 is the nominal noise covariance matrix and the norm is any valid matrix norm.

Proposition 7

The least favorable noise in N_1 is

$$n_L = n_0 + \varepsilon I \tag{3.43}$$

Proof

By means of Lemma 3.2, this is equivalent to proof that n_L is a maximal element of N_1. For any $n \in N_1$ and $x \in \mathbb{R}^k$:

$$x^T(n_L - n)x = x^T(n_0 + \varepsilon I - n)x = \varepsilon\|x\|^2 + x^T(n_0 - n)x \ ; \tag{3.44}$$

but by the Schwarz inequality

$$|x^T(n_0 - n)x| \leq \|x\| \ \|(n_0 - n)x\| \tag{3.45}$$

$$\leq \|x\|^2 \|n_0 - n\| \ , \tag{3.46}$$

where the last inequality must hold for any type of matrix norm [7]. Combining (3.44) and (3.46) we get

$$x^T(n_L - n)x \geq \|x\|^2(\varepsilon - \|n_0 - n\|) \geq 0 \tag{3.47}$$

Finally, it is easy to see that $n_L \in N_1$ since for any $x \in \mathbb{R}^k$, $x \neq 0$:

$$x^T n_L x = x^T n_0 x + \varepsilon\|x\|^2 > 0 \tag{3.48}$$

Note that the expression $\langle h, nh \rangle$ can be put as the inner product of matrices $[n, hh^T]$ defined by

$$[a,b] = \mathrm{tr}\,\{a \cdot b^T\} \tag{3.49}$$

Then other uncertainty classes can be solved analytically via this inner product. The reader is referred to Section 4.5.

3.2.3 Uncertainty in Signal and Noise

In this section we suppose that we receive some signal $s \in S$ and some noise with covariance $n \in N$. We must seek for a least favorable pair (s_L, n_L) that fulfills conditions 2 and 3 of Theorem 3.2. If S and N are such that analytical solutions for the least favorable signal and noise are available, we get a set of two equations in two unknowns: s_L and n_L, whose solution exists if and only if there exists a *least favorable pair* in (S,N).

Further simplification is obtained when one or both of the equations gives s_L and/or n_L directly, i.e. s_L (n_L) does not depend on the input noise (signal). As an example consider the case in which $S = S_1$ (3.16) and $N = N_1$ (3.42). Recall that the least favorable signal and noise were given respectively by:

$$s_L = s_0 - \sigma^2 h_L$$

$$n_L = n_0 + \varepsilon I$$

Therefore the robust matched filter $h_L = n_L^{-1} s_L$ is

$$h_L = (n_0 + (\varepsilon + \sigma^2) I)^{-1} s_0 \tag{3.50}$$

with $\sigma^2 \| h_L \| = \Delta$. Note that if the nominal noise is white, the nominal matched filter $(h_0 = s_0)$ is robust for uncertainties in signal and noise defined by S_1 and N_1 respectively.

3.2.4 Signal Selection

In this section we address the point of optimum signal design for matched filtering in the presence of uncertainties. We will assume (as is generally the case in signal selection) that the power of the signal to be selected is constrained. When the incoming signal and noise are known and are processed by their matched filter, it is well known that the signal that maximizes the SNR is the minimum-eigenvalue eigenvector of the noise covariance matrix. However, as we will see, when the received signal is deviated from the one sent this is, in general, no longer true.

In essence, the problem of signal selection under uncertainties is

$$s_m = \arg \max_{s_0, \|s_0\| = c} \; \max_{h \in \mathcal{H}} \; \inf_{(s,n) \in \mathcal{S}(s_0) \times \mathcal{N}} \rho(h;s,n) \tag{3.51}$$

where in order to simplify matters we have assumed that the noise and signal uncertainty classes are independent, and that the signal uncertainty class is the deviation from a given nominal (like the models treated in 3.2.2). It is clear that if the signal is known, it has to be designed to give optimal SNR with the least favorable noise, i.e. it must be its minimum-eigenvalue eigenvector. If the noise is known, we investigate the signal selection problem for the same uncertainty classes that we did in 3.2.2. In any case (3.51) results in

$$\begin{aligned} s_m &= \arg \max_{s_0, \|s_0\| = c} \; \min_{s \in \mathcal{S}(s_0)} \langle s, n^{-1} s \rangle \\ &= \arg \max_{s_0, \|s_0\| = c} \langle h_L(s_0), n h_L(s_0) \rangle \; . \end{aligned} \tag{3.52}$$

i) $\underline{\ell_2 \text{ uncertainty}}$. Recalling (3.22), we have

$$\langle h_L(s_0), n h_L(s_0)\rangle = s_0^T(n+\sigma^2 I)^{-1} n\ (n+\sigma^2 I)^{-1} s_0 \ . \tag{3.53}$$

It is easy to prove that x is eigenvector of n with eigenvalue λ, if and only if x is eigenvector of $(n+\sigma^2 I)^{-1} n\ (n+\sigma^2 I)^{-1}$ with eigenvalue $\lambda/(\lambda+\sigma^2)^2$. Then, one tends to think that the optimum s_0 is the minimum-eigenvalue eigenvector of $(n+\sigma^2 I)^{-1} n\ (n+\sigma^2 I)^{-1}$ (and therefore an eigenvector of n); however, this is not necessarily the case, because σ depends on s_0. Let $\{\varphi_1,\ldots,\varphi_k\}$ be a family of orthonormal eigenvectors of n. Then with $s_0 = \sum_{i=1}^{k} a_i\varphi_i$

$$s_0^T(n+\sigma^2 I)^{-1} n\ (n+\sigma^2 I)^{-1} s_0 = \sum_{i=1}^{k} \lambda_i a_i^2/(\lambda_i+\sigma^2)^2 \tag{3.54}$$

$$\|h_L(s_0)\| = \Big(\sum_{i=1}^{k} a_i^2/(\lambda_i+\sigma^2)^2\Big)^{\frac{1}{2}} \tag{3.55}$$

So (3.52) results in the following nonlinear optimization problem:

$$\arg\max_{\{a_i\}} \sum_{i=1}^{k} \lambda_i a_i^2/(\lambda_i+\sigma^2)^2 \tag{3.56}$$

subject to

$$\sigma^4 \sum_{i=1}^{k} a_i^2/(\lambda_i+\sigma^2)^2 = \Delta^2 \tag{3.57}$$

$$\sum_{i=1}^{k} a_i^2 = c^2 \tag{3.58}$$

ii) $\underline{\ell_\infty \text{ uncertainty}}$. When the noise is uncorrelated we obtained the explicit solution for the least favorable signal in (3.27). Applying this expression to (3.52) we get

$$s_m = \arg \max_{s_0, \| s_0 \| = c} SM(s_0)$$

$$SM(s_0) = \langle s_L(s_0), n^{-1} s_L(s_0) \rangle$$

$$= \sum_{s_{0i} > \Delta} \frac{1}{\lambda_i} (s_{0i} - \Delta)^2 + \sum_{s_{0i} < -\Delta} \frac{1}{\lambda_i} (s_{0i} + \Delta)^2 \tag{3.59}$$

where $\{\lambda_i\}$ is the set of eigenvalues of n.

Proposition 8

The optimal nominal signal for robust matched filtering when the noise has a known diagonal covariance matrix and the uncertainty on the received signal is modeled by an ℓ_∞ class, is a minimum-eigenvalue eigenvector of the noise covariance.

Proof

We suppose implicitly that the signal uncertainty is small compared with its power (specifically $k\Delta^2 < c^2$). First we show that if s_0 is such that, for some $i, 0 < |s_{0i}| \leq \Delta$, there exists s_0^* such that $\| s_0^* \| = \| s_0 \| = c$ and $SM(s_0^*) > SM(s_0)$. Let

$$s_{0j}^* = \begin{cases} 0 & j = i \\ s_{0j} & j \neq i,m \\ \sqrt{s_{0i}^2 + s_{0m}^2} & j = m \end{cases} \tag{3.60}$$

with $s_{0m} = \max_i \{|s_{0i}|\} > \Delta$

$$SM(s_0^*) - SM(s_0) = \frac{1}{\lambda_m} [(\sqrt{s_{0i}^2 + s_{0m}^2} - \Delta^2)^2 - (|s_{0i}| - \Delta)^2] > 0 \tag{3.61}$$

Also, if s_0 is such that $s_{0i} > \Delta$ with $\lambda_i > \lambda_p = \min \{\lambda_j\}$, there exists s_0^* such that $\| s_0^* \| = \| s_0 \| = c$ and $SM(s_0^*) > SM(s_0)$. Let s_0^* be defined by (3.60) replacing m by p.

$$SM(s_0^*) - SM(s_0) = \frac{1}{\lambda_p} [(|s_{0p}^*| - \Delta)^2 - (|s_{0p}| - \Delta)^2] - \frac{1}{\lambda_i} (|s_{0i}| - \Delta)^2$$

$$> \frac{1}{\lambda_p} [2\Delta(|s_{0p}| + |s_{0i}| - |s_{0p}^*|) - \Delta^2] \; . \tag{3.62}$$

Now consider

$$(|s_{0p}| + |s_{0i}| - \frac{\Delta}{2})^2 - |s_{0p}^*|^2 = 2|s_{0i}||s_{0p}| - \Delta(|s_{0p}| + |s_{0i}|) + \frac{\Delta^2}{4}$$

$$= (\varepsilon + \delta)\Delta + 2\varepsilon\delta > 0 \; , \tag{3.63}$$

where $\varepsilon = |s_{0i}| - \Delta$, $\delta = |s_{0p}| - \Delta$ are positive by supposition (otherwise (3.61) shows already that s_0^* is not optimal). Since $|s_{0p}| + |s_{0i}| > \frac{\Delta}{2}$, (3.63) implies that the right side of (3.62) is positive, and thus we get the desired result. Therefore, we have shown that the only s_0 for which there does not exist a different signal with better SM is the one having all the samples corresponding to eigenvalues greater than the minimum, equal to zero.

4. LINEAR OBSERVERS AND REGULATORS

4.1 Introduction

Undoubtedly, modern control is one of the areas in which the minimax approach to problems with uncertain models has flourished more markedly. Specifically, the great practical impact of Kalman filtering has resulted in a sizable portion of technical literature related to the design of optimal observers for systems whose noise statistics are not accurately known. However, up to now, the problem has not been treated with full generality; the available results are related either to the steady-state case or to particular classes of uncertainties. Moreover, the main relevance of this application of our general results for minimax robust filtering, stems from the fact that the payoff functions that we encounter in these filtering situations are not concave-convex and thus the use of the conventional theorems of game theory is not possible (we will show that the often quoted treatment of D'Appolito and Hutchinson [19] that relies on those results, is incorrect).

As in the other cases, two kinds of results will be presented, first a minimax theorem, that gives a saddle point solution to the corresponding game by showing that the minimax equality holds, and thus the solution is the optimal filter for the least favorable operating point, and second a procedure to find these least favorable situations for general uncertainty classes. In order not to be overly repetitive, our emphasis will be in the discrete-time case, presenting the main results in continuous-time as well. We will analyze the following filtering situations:

i) the linear observer problem for predictor state estimation ($\hat{x}_{k/k-1}$)

ii) the linear observer problem for filter state estimation ($\hat{x}_{k/k}$)

iii) the regulator problem for linear quadratic optimal control.

Finally, several discrete-time examples will illustrate the search for least favorables for general uncertainty classes of covariance matrices of practical interest.

For convenience, we repeat in the form of lemmas the results from Chapter 2 that will be used in this chapter.

<u>Lemma 1</u>

Suppose that $\delta(h,\cdot)$ is concave in Q for every $h \in \mathcal{H}^*$.

If a least favorable operating point -- for $(Q,\mathcal{H}^*,\delta)$ -- q_L, and its optimum filter h_L, form a regular pair, then h_L is a robust filter for $(P,\mathcal{H},\delta)$.

<u>Lemma 2</u>

Suppose that Q is a cartesian product of sets, $Q = Q_1 \times \dots \times Q_k$ and that the penalty function can be put as

$$\delta(h,(q_1,\dots,q_k)) = F(f_1(h,q_1),\dots,f_k(h,q_k)) ,$$

with F nondecreasing in each one of its arguments. Denote, for $i = 1,\dots,k$

$$q_i^*(h) = \arg \max_{q_i \in Q_i} f_i(h,q) .$$

If h_L is a solution of the equation

$$h_L = \arg \min_{h \in \mathcal{H}^*} \delta(h,(q_1^*(h_L),\dots,q_k^*(h_L)) , \tag{2.20}$$

then h_L is a robust filter for $(P,\mathcal{H},\delta)$.

Lemma 3

Under the suppositions of Lemmas 1 and 2, $q_L = (q_1^*(h_L),\ldots,q_k^*(h_L))$ is a least favorable operating point -- for $(Q,\mathcal{H}^*,\delta)$ -- if and only if h_L is a solution to (2.20).

4.2 Robust observer problem

Consider the following linear discrete-time system

$$x_{k+1} = A_k x_k + B_k u_k + w_k + \bar{w}_k, \quad k_o \leq k \leq N-1 \tag{4.1}$$

$$z_k = C_k x_k + v_k + \bar{v}_k, \quad k_o \leq k \leq N-1 \tag{4.2}$$

with x_k an $n \times 1$ state vector, u_k an $m \times 1$ control vector and z_k an $r \times 1$ output vector. The initial state (x_{k_o}) is a random vector with mean m_o and variance Σ_o, and $(w_k + \bar{w}_k)$, $(v_k + \bar{v}_k)$ are random sequences independent of the initial state, representing the process and observation noises with means $\bar{w}_k$ and $\bar{v}_k$ and covariance

$$\operatorname{cov}\left(\begin{bmatrix} w_k \\ v_k \end{bmatrix}, \begin{bmatrix} w_\ell \\ v_\ell \end{bmatrix}\right) = \begin{bmatrix} \Xi_k & \Psi_k \\ \Psi_k^T & \Theta_k \end{bmatrix} \delta_{k\ell} \tag{4.3}$$

We suppose that the means and the covariance matrices are known only to belong to some uncertainty classes, $\operatorname{col}(m_o, \bar{w}_{k_o}, \ldots, \bar{w}_{N-2}, \bar{v}_{k_o}, \ldots, \bar{v}_{N-2}) = y \in \mathcal{M}$, $(\Xi_k, \Psi_k, \Theta_k, \Sigma_o)_{k=k_o}^{N-1} = \Lambda \in \mathcal{C}$, such that Ξ_k and Σ_o are positive semidefinite, Θ_k is positive definite and Ψ_k is congruent with Ξ_k and Θ_k. In the presence of these uncertainties the optimal observer and regulator problems will be solved in a minimax sense, i.e., our goal is to find

$$h_R = \arg \min_{h \in \mathcal{H}} \sup_{(y,\Lambda) \in \mathcal{M} \times \mathcal{C}} \delta(h,(y,\Lambda)) \tag{4.4}$$

where $\mathcal{H}$ is some class of allowable filters and $\delta(\cdot,\cdot)$ is a penalty function. In order to solve this game $(M \times C, \mathcal{H}, \delta)$ we will apply the general results of Chapter 2.

1. Formulation

The optimal linear observer of the state of (4.1) is the solution to

$$\arg \min_{H \in \mathcal{H}} E[e^T_{k/q} Q_k e_{k/q} | U_q, Z_q], \tag{4.5}$$

where $\mathcal{H}$ is the set of linear filters of appropriate dimensionality with state $\hat{x}_{k/q}$, $e_{k/q} \triangleq x_k - \hat{x}_{k/q}$, $U_q = \{u_{k_o}, \ldots, u_q\}$, $Z_q = \{z_{k_o}, \ldots, z_q\}$ and Q_k is a positive semidefinite weighting matrix. When the second order statistics of the stochastic system (4.1-3) are known, the solution for every time k is given by the Kalman filter; the cases of most general interest, i.e., $q = k-1$, (one-step) predictor estimate, and $q = k$, filter estimate, are given by (e.g. [15]) recursive formulas of the form

$$\hat{x}_{k+1/k} = A_k \hat{x}_{k/k-1} + B_k u_k + \bar{w}_{Lk} + K_k(z_k - \bar{v}_{Lk} - C_k \hat{x}_{k/k-1}), \tag{4.6}$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + E_k(z_k - \bar{v}_{Lk} - C_k \hat{x}_{k/k-1}), \tag{4.7}$$

$$\hat{x}_{k_o/k_o-1} = m_{Lo}. \tag{4.8}$$

Subtracting (4.6) from (4.1) and x_k from both sides of (4.7) we get the error dynamical equations for arbitrary K_k, E_k, m_{Lo}, $\bar{w}_{Lk}$, $\bar{v}_{Lk}$:

$$e_{k+1/k} = (A_k - K_k C_k) e_{k/k-1} - K_k v_k + w_k + \bar{w}_k - \bar{w}_{Lk} - K_k(\bar{v}_k - \bar{v}_{Lk}), \tag{4.9}$$

$$e_{k/k} = (I - E_k C_k) e_{k/k-1} - E_k v_k - E_k(\bar{v}_k - \bar{v}_{Lk}), \tag{4.10}$$

$$e_{k_o/k_o-1} = x_{k_o} - m_{Lo}. \tag{4.11}$$

Denoting the error variances by $\Sigma^a_{k/q} = \mathrm{var}\,[e_{k/q}]$, the vector of means in the estimator by y_L (defined analogously to y), and the state-transition of the fundamental matrix $(A - KC)$ by $\Phi(\cdot,\cdot)$ and taking into account that the white sequences w_k, v_k are independent of $e_{k/k-1}$, (4.9-11) lead to

$$E\,[e_{k/k-1}\, e^T_{k/k-1}] = \Sigma^a_{k/k-1} + \Phi_k(IK)(y-y_L)(y-y_L)^T(IK)^T\Phi^T_k\,, \tag{4.12}$$

$$\Sigma^a_{k/k-1} = \Phi(k,k_o)\Sigma_o\,\Phi^T(k,k_o) + \sum_{j=k_o}^{k-1}\Phi(k,j+1)\,M_j\Phi^T(k,j+1)\,, \tag{4.13}$$

$$M_k = [\,I \,:\, -K_k]\begin{bmatrix}\Xi_k & \Psi_k \\ \Psi^T_k & \Theta_k\end{bmatrix}\begin{bmatrix} I \\ \hline -K^T_k\end{bmatrix}\,, \tag{4.14}$$

$$\Phi_k = [\,\Phi(k,k_o) \,:\, \dots \,:\, \Phi(k,k) \,:\, 0_{n\times(N-1-k)n}]\,, \tag{4.15}$$

$$(IK) = \left[\begin{array}{c|c} & 0_{n\times(N-1-k_o)r} \\ \hline I_{(N-k_o)n} & \begin{array}{c|c} -K_{k_o} & \\ \hline & \ddots \\ & -K_{N-2} \end{array} \end{array}\right]; \tag{4.16}$$

$$E\,[e_{k/k}\, e^T_{k/k}] = \Sigma^a_{k/k} + F_k(y-y_L)(y-y_L)^T\,F^T_k\,, \tag{4.17}$$

$$\Sigma^a_{k/k} = [\,I - E_kC_k]\,\Sigma^a_{k/k-1}\,[\,I - E_kC_k]^T + E_k\Theta_kE^T_k\,, \tag{4.18}$$

$$F_k = [\,I - E_kC_k]\,\Theta_k\,(IK) - E_k\,[\,0_{(N-1-K_o)r\times(N-K_o)n} \,\vdots\, I_{(N-1-K_o)r}]\,. \tag{4.19}$$

Here the subindices of the null and identity matrices indicate their respective size.

2. Uncertainty in Covariances

We suppose now that the means are known, i.e. $y = y_L$, and therefore that the estimates are unbiased. By the definition of the state-transition matrix it is easy to check [16] that $\Sigma^a_{k/k-1}$ satisfies the equation

$$\Sigma^a_{k+1/k} = [A_k - K_k C_k]\, \Sigma^a_{k/k-1}\, [A_k - K_k C_k]^T + M_k \quad . \tag{4.20}$$

The optimal gains are given by

$$K^*_k = \arg \min_{K_k \in \mathbb{R}^{n \times r}} \Sigma^a_{k/k-1}$$

$$= [A_k\, \Sigma_{k/k-1}\, C^T_k + \Psi_k]\,[C_k\, \Sigma_{k/k-1}\, C^T_k + \Theta_k]^{-1}, \tag{4.21}$$

$$\Sigma_{k+1/k} = A_k\, \Sigma_{k/k-1}\, A^T_k + \Xi_k -$$

$$[A_k\, \Sigma_{k/k-1}\, C^T_k + \Psi_k]\,[C_k\, \Sigma_{k/k-1}\, C^T_k + \Theta_k]^{-1}[A_k\, \Sigma_{k/k-1}\, C^T_k + \Psi_k]^T, \tag{4.22}$$

$$\Sigma_{k_o/k_o-1} = \Sigma_o\,,$$

$$E^*_k = \arg \min_{E_k \in \mathbb{R}^{n \times r}} \Sigma^a_{k/k} = \Sigma_{k/k-1}\, C^T_k\, [C_k\, \Sigma_{k/k-1}\, C^T_k + \Theta_k]^{-1}, \tag{4.23}$$

where

$$\Sigma_{k/q} = \min_{K_k, E_k \in \mathbb{R}^{n \times r}} \Sigma^a_{k/q} \tag{4.24}$$

can be precomputed since it does not depend on the observations.

Since all the matrices involved are possibly time-varying, in the remainder of the discussion we will drop, for convenience, their explicit dependence on time except when this could be ambiguous.

In order to prove the minimax theorems we will make use of the following sensitivity result.

<u>Lemma 4</u>

Suppose C' is a convex set and $(\Xi, \Psi, \Theta, \Sigma_o) = \Lambda = (1-\alpha)\,\Lambda_L + \alpha\Omega$ with $\Lambda_L = (\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ and $\Omega = (V, X, Y, Z)$ both belonging to C'. Let $\hat{x}^L_{k/q}$ be the state estimate of the Kalman filter designed for Λ_L when Λ is truly present. With $\bar{e}_{k/q} \triangleq x_k - \hat{x}^L_{k/q}$ and $\Gamma_{k/q} \triangleq E[\bar{e}_{k/q}\, \bar{e}^T_{k/q}]$, we have that for all $\Lambda_L, \Omega \in C'$ and time k, and for $q = k, k-1$

$$\lim_{\alpha \to o} \frac{1}{\alpha} (\Sigma_{k/q} - \Gamma_{k/q}) = 0 \tag{4.25}$$

<u>Proof</u>

$\Sigma_{k/k-1}$ and $\Gamma_{k/k-1}$ satisfy the equations (4.14) and (4.20) for the optimal Kalman gains K^* and K^*_L designed for Λ and Λ_L respectively. Note that in both cases the same set of covariances $(\Xi, \Psi, \Theta, \Sigma_o)$ is present. Therefore, we have

$$\begin{aligned}
\Sigma_{k+1/k} - \Gamma_{k+1/k} &= [A - K^*_L C]\,(\Sigma_{k/k-1} - \Gamma_{k/k-1})\,[A - K^*_L C]^T \\
&\quad - \Psi (K^* - K^*_L)^T - (K^* - K^*_L)\,\Psi^T + K^* \Theta K^{*T} - K^*_L \Theta K^{*T}_L \\
&\quad - (K^* - K^*_L)\, C\, \Sigma_{k/k-1}\, A^T - A\, \Sigma_{k/k-1}\, C^T (K^* - K^*_L)^T \\
&\quad - K^*_L C\, \Sigma_{k/k-1}\, C^T K^{*T}_L + K^* C\, \Sigma_{k/k-1}\, C^T K^{*T} \\
&= [A - K^*_L C]\,(\Sigma_{k/k-1} - \Gamma_{k/k-1})\,[A - K^*_L C]^T \\
&\quad - (K^* - K^*_L)\,[C\, \Sigma_{k/k-1}\, C^T + \Theta]\,(K^* - K^*_L)^T
\end{aligned} \tag{4.26}$$

where in order to get the last equality the expression for K^* (4.21) has been taken into account. It is worth noting that (4.26) gives the variation of the error variance due to an arbitrary modification (up to here the value of K^*_L has not been used) of the optimal Kalman gain (cf. [17]). Because $\lim_{\alpha \to o} \frac{1}{\alpha} [K^* - K^*_L]$ exists for every time k and $\Lambda_L, \Omega \in C'$ (see Appendix 1), in the limit (4.26) reduces to a homogeneous equation, and since $\Sigma_{k_o/k_o-1} = \Sigma_o = \Gamma_{k_o/k_o-1}$ we have

$$\lim_{\alpha \to o} \frac{1}{\alpha} (\Sigma_{k/k-1} - \Gamma_{k/k-1}) = 0 \tag{4.27}$$

Using (4.12) for E^* and E_L^*, again the gains designed for Λ and Λ_L respectively, and the expression (4.23) for E^*, a series of manipulations result in

$$\begin{aligned}\Sigma_{k/k} - \Gamma_{k/k} &= [I - E_L^* C]\,(\Sigma_{k/k-1} - \Gamma_{k/k-1})\,[I - E_L^* C]^T \\ &\quad - (E^* - E_L^*)\,[C\,\Sigma_{k/k-1}\,C^T + \Theta]\,(E^* - E_L^*)^T\end{aligned} \tag{4.28}$$

Now, taking $\lim_{\alpha \to o} \frac{1}{\alpha}[\cdot]$ of both sides of (4.28), the first term of the right side is zero as a consequence of (4.27) and the second term vanishes as well, because $\lim_{\alpha \to o} \frac{1}{\alpha}(E^* - E_L^*)$ exists as can be shown similarly to Appendix 1. This ends the proof of the Lemma.

If when the second order statistics are uncertain, we use the weighted square error (4.5) as the penalty function, inasmuch as the history of the least favorable statistics can be different for every k, it is possible that the minimax filter is given by a recursion that does not solve the problem for previous times, therefore we are forced to build a different filter for every k, running from k_o to k. The inapplicability of this solution leads to the consideration of the penalty function (cf. [13])

$$L(H,y,\Lambda) = \sum_{k=k_o}^{N-1} E\,[e_{k/q}^T\, Q_k\, e_{k/q} | U_q, Z_q] \tag{4.29}$$

that when the statistical model is known has the same minimizing filter as (4.5). Furthermore, note that the penalty function of (4.5) can be considered a particular case of (4.29).

We assume that C is an uncertainty class such that there exists a convex set $C' \supset C$ that fulfills (2.13) achieving the inf sup, i.e. with least favorables. Given that the means are known and $\mathcal{H}$ is the set of linear filters, the following minimax theorem holds.

Theorem 1

The Kalman filter for the least favorable collection of covariances for (predictor/filter) state estimation is robust for the game $(\mathcal{C},\mathcal{H},L)$.

Proof

First consider the class $\mathcal{H}^*$ of linear filters of the form (4.6-8) for $K,E \in \mathbb{R}^{n \times r}$ and with matrices A,B,C given by the original system (4.1-2). Then the penalty function is

$$L(H,\Lambda) = \sum_{k=k_o}^{N-1} \mathrm{tr}\{Q_k \Sigma^a_{k/q}\} \tag{4.30}$$

with $\Sigma^a_{k/q}$ given by (4.13), (4.14), (4.18) for $q = k,k-1$. Now we check that the sufficient conditions of Lemma 1 hold in this case: i) $L(H,\Lambda)$ is linear in Ξ, Ψ, Θ, Σ_o and thus concave in C'; ii) it follows from Lemma 4 that every element in C' forms a regular pair with its associated Kalman filter; and iii) since the Kalman filter (optimum over $\mathcal{H}$) for a given Λ, belongs always to the class $\mathcal{H}^*$ we have condition ii) of Theorem 2.1 for any set of covariances. (Note that if x_{k_o} is a gaussian random vector and the process and observation noises are jointly gaussian, we can drop the restriction of linearity for the class $\mathcal{H}$ in the Theorem.) Thus the assumptions of Lemma 1 are satisfied and Theorem 1 follows.

Remark

The last proof allows us to notice the role of the class $\mathcal{H}^*$ in Th. 2.1 Thanks to the previous restriction to the set of filters with given A, B, C matrices we were able to prove the concavity in the uncertainties of the penalty function, while the theorem holds for the general class of filters.

(This restriction of the set of filters has been previously justified by the convenience of an unbiased estimate [13]. However, this is not enough to prove the optimality of the solution over all possible linear estimators.)

Theorem 1 has reduced the original minimax filter design problem to the search for least favorable sets of covariances. In this point the application of Lemma 3 has proven to be successful in dealing with important uncertainty classes in other contexts [5]. In the present case, let us define

$$W_j(R) \triangleq \sum_{k=j}^{N-1} \Phi_L^T(k,j)\, R_k\, \Phi_L(k,j) \tag{4.31}$$

where $R_k \in \mathbb{R}^{n\times n}$ (if $R > 0$ then $W(R) > 0$), and $\Phi_L(\cdot,\cdot)$ is the state transition matrix of the fundamental matrix $(A - K_L C)$. Using this definition and interchanging the order of summation, (4.13) and (4.18) result in

$$\sum_{k=k_o}^{N-1} \mathrm{tr}\{Q_k\, \Sigma^a_{k/k-1}\} = \mathrm{tr}\{\Sigma_o\, W_{k_o}(Q) + \sum_{j=k_o}^{N-2} M_j\, W_{j+1}(Q)\} \tag{4.32}$$

$$\begin{aligned}\sum_{k=k_o}^{N-1} \mathrm{tr}\{Q_k\, \Sigma^a_{k/k}\} = \mathrm{tr}\{&\Sigma_o\, W_{k_o}([I - E_L C]^T Q[I - E_L C]) \\ &+ \sum_{j=k_o}^{N-2} M_j\, W_{j+1}([I - E_L C]^T Q[I - E_L C]) \\ &+ \sum_{k=k_o}^{N-1} E_{Lk}^T\, Q_k\, E_{Lk}\, \Theta_k\}\end{aligned} \tag{4.33}$$

Applying Lemma 3 to these expressions, Theorems 2 and 3 follow straightforwardly.

Theorem 2

If C' can be put as the cartesian product of uncertainty classes $\chi \times P \times \mathcal{N} \times S$, then $(\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ is the least favorable collection of covariances for predictor state estimation if and only if the following are satisfied

$$\Xi_L = \arg \max_{\Xi \in \chi} \operatorname{tr}\{ \sum_{j=k_o}^{N-2} \Xi_j W_{j+1}(Q)\} \tag{4.34}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr}\{ \sum_{j=k_o}^{N-2} \Psi_j K_{Lj}^T W_{j+1}(Q)\} \tag{4.35}$$

$$\Theta_L = \arg \max_{\Theta \in \mathcal{N}} \operatorname{tr}\{ \sum_{j=k_o}^{N-2} \Theta_j K_{Lj}^T W_{j+1}(Q) K_{Lj}\} \tag{4.36}$$

$$\Sigma_{oL} = \arg \max_{\Sigma_o \in S} \operatorname{tr}\{\Sigma_o W_{k_o}(Q)\} \tag{4.37}$$

$$K_{LK} = [A_k \Sigma^L_{k/k-1} C_k^T + \Psi_{Lk}] [C_k \Sigma^L_{k/k-1} C_k^T + \Theta_{Lk}]^{-1} \tag{4.38}$$

$$\Sigma^L_{k+1/k} = A_k \Sigma^L_{k/k-1} A_k^T + \Xi_{Lk} - [A_k \Sigma^L_{k/k-1} C_k^T + \Psi_{Lk}][C_k \Sigma^L_{k/k-1} C_k^T + \Theta_{Lk}]^{-1}[A_k \Sigma^L_{k/k-1} C_k^T + \Psi_{Lk}]^T \tag{4.39}$$

$$\Sigma^L_{k_o/k_o-1} = \Sigma_{oL} \tag{4.40}$$

Notice that while the optimal observer for fixed covariances is independent of the error weighting matrix Q, this is not the case here, since the least favorable covariances may depend on Q. Moreover, the least favorable covariances (from $k = k_o$ to $N-1$) depend on the final time N.

Theorem 3

Under the hypothesis of Theorem 2, $(\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ is the least favorable collection of covariances for filter state estimation if and only if (4.38-40) and the following are satisfied

$$\Xi_L = \arg \max_{\Xi \in X} \operatorname{tr}\{ \sum_{j=k_o}^{N-2} \Xi_j \, P_{j+1} \} \tag{4.41}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr}\{ \sum_{j=k_o}^{N-2} \Psi_j \, K_{Lj}^T \, P_{j+1} \} \tag{4.42}$$

$$\Theta_L = \arg \max_{\Theta \in N} \operatorname{tr}\{ \sum_{j=k_o}^{N-1} \Theta_j \, E_{Lj}^T \, Q_j \, E_{Lj} + \sum_{j=k_o}^{N-2} \Theta_j \, K_{Lj}^T \, P_{j+1} \, K_{Lj} \} \tag{4.43}$$

$$\Sigma_{oL} = \arg \max_{\Sigma_o \in S} \operatorname{tr}\{\Sigma_o \, P_{k_o} \} \tag{4.44}$$

$$P_j = W_j([I - E_L C]^T \, Q \, [I - E_L C]) \tag{4.45}$$

$$E_{Lk} = \Sigma_{k/k-1}^L \, C_k^T \, [C_k \, \Sigma_{k/k-1}^L \, C_k^T + \Theta_{Lk}]^{-1} \tag{4.46}$$

3. Uncertainty in Means and Covariances

Once we have solved the problem for uncertainty in the covariance matrices, we drop the assumption that the means are known and we attack the completely general question. The penalty functions for predictor and filter state estimation are derived from (4.5), (4.12), (4.17)

$$L_p(h,y,\Lambda) = \sum_{k=k_o}^{N-1} (\operatorname{tr}\{Q_k \, \Sigma_{k/k-1}^a\} + (y-y_L)^T (IK)^T \, \Phi_k^T \, Q_k \, \Phi_k (IK)(y-y_L)) \tag{4.47}$$

$$L_f(h,y,\Lambda) = \sum_{k=k_o}^{N-1} (\operatorname{tr}\{Q_k \, \Sigma_{k/k}^a\} + (y-y_L)^T \, F_k^T \, Q_k \, F_k(y-y_L)) \tag{4.48}$$

It is not possible to apply Lemmas 1 and 3 to these penalty functions because they are not concave in the uncertainties (in fact, they are convex). However, we can put the uncertainty set as the cartesian product $\mathcal{X} \times \mathcal{P} \times \mathcal{N} \times \mathcal{S} \times \mathcal{M}$, and the part of the penalty function due to the error variance can be decomposed in the same way as in Theorems 2 and 3 in order to apply Lemma 2 (note that no further decomposition of the penalty due to the means is possible). Unfortunately, except for trivial mean uncertainty classes, equation (2.20) has no solution in this case. The reason for this can be seen intuitively considering that the performance of a filter built for a given set of means can only be deteriorated when different means are truly present, and hence no saddle-point solution to the filtering game exists.

Alternatively, we deal with the <u>soft minimax</u> solution to the problem. According to this approach, useful to model frequent situations in which there is not "total" uncertainty, operating points that are closer to a given nominal are more likely to occur than those in the uncertainty class more distant from it. This further knowledge can be taken advantage of, by adding to the penalty function an additional term accounting for the distance between the operating point and the nominal. Applying this to the uncertainty in means

$$L'_p(h,y,\Lambda) = L_p(h,y,\Lambda) - (y - y_N)^T D (y - y_N) \tag{4.49}$$

where D is a positive semidefinite matrix, and y_N is the vector of nominal means.

<u>Theorem 4</u>

Suppose for all possible covariances we have

$$\sum_{k=k_o}^{N-1} (IK)^T \Phi_k^T Q_k \Phi_k (IK) - D \leq 0 \tag{4.50}$$

then the Kalman filter given by (4.34-40) and with nominal means is robust for the game $(\mathcal{M} \times \mathcal{C}, \mathcal{K}, L'_p)$ (predictor state estimation).

Proof

Applying Lemma 2, we see that in order for (2.20) to have a solution, the Kalman gain must be the optimal for the least favorable set of covariances (independent of the means) and thus (4.34-37) must hold. Besides we need a solution to

$$y_R = \arg \min_{y_L \in \mathcal{M}} (y^*(y_R) - y_L)^T \sum_{k=k_o}^{N-1} (IK)^T \Phi_k^T Q_k \Phi_k (IK)(y^*(y_R) - y_L)$$
$$- (y^*(y_R) - y_N)^T D (y^*(y_R) - y_N) \tag{4.51}$$

$$y^*(y_R) = \arg \max_{y \in \mathcal{M}} (y - y_R)^T \sum_{k=k_o}^{N-1} (IK)^T \Phi_k^T Q_k \Phi_k (IK)(y - y_R)$$
$$- (y - y_N)^T D (y - y_N) \tag{4.52}$$

But taking into account that,

$$0 \leq \sum_{k=k_o}^{N-1} (IK)^T \Phi_k^T Q_k \Phi_k (IK) \leq D \tag{4.53}$$

$y^*(y_N) = y_N$, and therefore y_N is solution to (4.51). An analogous proof holds for the next theorem.

Theorem 5

Suppose for all possible covariances we have

$$\sum_{k=k_o}^{N-1} F_k^T Q_k F_k - D < 0 \tag{4.54}$$

then the Kalman filter given by (4.38-46) and with nominal means is robust for the game $(\mathcal{M} \times C, \mathcal{K}, L_f^{'})$ (filter state estimation).

It is interesting to remark that the design of observers for systems with unknown input U_k is just a particular case of the unknown means problem treated here. Basar and Mintz [20] have solved a tracking-evasion problem through the design of a filter-estimator for a system with observation noise and initial state with known second order statistics and unknown input. They use a penalty function that counterbalances the mean square estimation error (in one instant) with a quadratic cost on the choice of the input. In our notation we have: for $k = k_o, \ldots, N-2$, $\bar{w}_k$ unknown, $\bar{w}_{Nk} = 0$, $\bar{v}_k = 0$, $\Psi_k = 0$, $\Xi_k = 0$, $Q_k = 0$; $Q_{N-1} \neq 0$, $m_o = 0$, $D = \mathrm{diag}(D_\infty^1, D_o, D_\infty^2)$ where D_∞^1, D_∞^2 are matrices of appropriate dimensions and arbitrarily large positive eigenvalues (this models the certainty on m_o and $\bar{v}_k$). Then the sufficient condition of Theorem 5 in order for the robust estimator to have $\bar{w}_{Lk} = 0$, reduces to

$$[I-E_{N-1}\, C_{N-1}]^T [\Phi(N-1,k_o+1) \vdots \cdots \vdots \Phi(N-1,N-1)]^T$$
$$Q_{N-1} [\Phi(N-1,k_o+1) \vdots \cdots \vdots \Phi(N-1,N-1)] [I-E_{N-1}\, C_{N-1}] - D_o < 0 \tag{4.55}$$

which is precisely the sufficient condition derived in [20] in order for the minimax estimator to have $U_k = 0$.

4.3 Robust Regulator Problem

1. Formulation

The optimal regulator for linear quadratic optimal control is the solution to

$$\arg \min_{H \in \mathcal{H}} J(H,y,\Lambda) \tag{4.56}$$

with

$$J(H,y,\Lambda) = E[x_N^T F x_N | Z_{N-1}] + \sum_{k=k_o}^{N-1} E[x_k^T Q_{1k} x_k + U_k^T Q_{2k} U_k | Z_{k-1}] \tag{4.57}$$

where x_k is the state of the system (4.1-2) with zero mean process noise, $\mathcal{H}$ is the set of linear filters, with input U_{k-1} and Z_{k-1} and output U_k, and F and Q_{1k} positive semidefinite and Q_{2k} positive definite. The optimal filter when the noise covariances are fixed is [16]:

$$U_k = -G_k \hat{x}_{k/k-1} \tag{4.58}$$

$$G_k = [Q_{2k} + B_k^T S_{k+1} B_k]^{-1} B_k^T S_{k+1} A_k \tag{4.59}$$

$$S_k = A_k^T S_{k+1} A_k + Q_{1k} - A_k^T S_{k+1} B_k [Q_{2k} + B_k^T S_{k+1} B_k]^{-1} B_k^T S_{k+1} A_k \tag{4.60}$$

$$S_N = F \tag{4.61}$$

and $\hat{x}_{k/k-1}$ the optimal predictor estimate given by (4.6), (4.8), (4.21), (4.22). This is the consequence of the separation principle of stochastic control, that enunciates the optimality of the feedback of the state estimate with the same gain as that in the known state case. Since the feedback gain G does not depend on the statistics of the noise it should be expected, that when these are unknown, the optimal value of G is unchanged, and thus the control and state estimation problems can still be solved separately. We will prove this statement rigorously and will show that in general the state estimates are not given by a minimax observer.

We will use the following identity [18].

Lemma 5

If the process noise has zero mean

$$E[x_N^T F x_N + \sum_{k=k_o}^{N-1} x_k^T Q_{1k} x_k + U_k^T Q_{2k} U_k] =$$

$$= m_o^T S_{k_o} m_o + tr\{S_{k_o} \Sigma_o\} + \sum_{k=k_o}^{N-1} tr\{S_{k+1} \Xi_k\}$$

$$+ \sum_{k=k_o}^{N-1} E[(U_k + G_k x_k)^T [B_k^T S_{k+1} B_k + Q_{2k}] (U_k + G_k x_k)] \qquad (4.62)$$

with G and S given by (4.59-61).

2. Uncertainty in Covariances

Theorem 6

The regulator consisting of the feedback for complete state information, of the state estimates produced by the Kalman filter for the least favorable collection of covariances for linear quadratic optimal control is robust for the game $(C,\mathcal{K},J)$.

Proof

Similarly to the proof of Theorem 1, we first consider the class $\mathcal{K}^*$ of linear filters of the form (4.6), (4.58-61) for $K \in \mathbb{R}^{n\times r}$ and with matrices A, B, C given by the original system. Let us define the positive semidefinite matrix

$$N_k \triangleq A_k^T S_{k+1} B_k [Q_{2k} + B_k^T S_{k+1} B_k]^{-1} B_k^T S_{k+1} A_k \tag{4.63}$$

then the penalty function can be expressed (by Lemma 3 and (4.58) and taking into account that the means are known) as

$$J(H,\Lambda) = m_o^T S_{k_o} m_o + tr\{S_{k_o} \Sigma_o\} + \sum_{k=k_o}^{N-1} tr\{S_{k+1} \Xi_k\}$$
$$+ \sum_{k=k_o}^{N-1} tr\{N_k \Sigma^a_{k/k-1}\} \tag{4.64}$$

Again this is linear in Ξ, Ψ, Θ, Σ_o, and thus concave in C'. Besides Lemma 4 implies that every element in C' forms a regular pair -- for $(C',\mathcal{H}^*,J)$ -- with its optimum regulator. The optimal linear regulator (over $\mathcal{H}$) for any set of covariances belongs to the class $\mathcal{H}^*$, so we have (2.7) for all $\Lambda \in C$, and the theorem follows.

Note that, as the following theorem states, the least favorable set of covariances for this problem may be different from the one for state estimation, hence our previous assertion that the Kalman filter for the minimax regulator is not a minimax observer, in general. By the same kind of manipulations that led to (4.32), the penalty function can be expressed as

$$J(H,\Lambda) = m_o^T S_{k_o} m_o + tr\{S_{k_o} \Sigma_o + S_N \Xi_{N-1}\}$$
$$+ tr\{\Sigma_o W_{k_o}(N) + \sum_{j=k_o}^{N-2} (\Xi_j S_{j+1} + M_j W_{j+1}(N))\} \tag{4.65}$$

Lemma 3 can be applied to this expression resulting in

Theorem 7

Under the same assumptions and definitions as in Theorem 2, $(\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ is the least favorable collection of covariances for linear quadratic optimal control if and only if (4.38-40) and the following are satisfied.

$$\Xi_L = \arg \max_{\Xi \in \chi} \operatorname{tr}\{\Xi_{N-1} F + \sum_{j=k_o}^{N-2} \Xi_j (S_{j+1} + W_{j+1}(N))\} \tag{4.66}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr}\{\sum_{j=k_o}^{N-2} \Psi_j K_{Lj}^T W_{j+1}(N)\} \tag{4.67}$$

$$\Theta_L = \arg \max_{\Theta \in N} \operatorname{tr}\{\sum_{j=k_o}^{N-2} \Theta_j K_{Lj}^T W_{j+1}(N) K_{Lj}\} \tag{4.68}$$

$$\Sigma_{oL} = \arg \max_{\Sigma_o \in S} \operatorname{tr}\{\Sigma_o (S_{k_o} + W_{k_o}(N))\} \tag{4.69}$$

3. Uncertainty in Means and Covariances

When the means of the initial state and of the observation noise are unknown, the penalty function is, using soft minimax,

$$\begin{aligned} J(h,y,\Lambda) = {} & \operatorname{tr}\{S_{k_o} \Sigma_o\} + \sum_{k=k_o}^{N-1} \operatorname{tr}\{S_{k+1} \Xi_k\} + \sum_{k=k_o}^{N-1} \operatorname{tr}\{N_k \Sigma^a_{k/k-1}\} \\ & + (V - V_L)^T (JK)^T \sum_{k=k_o}^{N-1} \Phi_k^T N_k \Phi_k (JK)(V - V_L) + V^T (IS) V \\ & - (V - V_N)^T D (V - V_N) \end{aligned} \tag{4.70}$$

where $(JK) = \operatorname{diag}\{I_n, -K_{k_o}, \ldots, -K_{N-2}\}$; $(IS) = \operatorname{diag}\{S_{k_o}, 0_{(N-1-k_o)r}\}$, $V = \operatorname{col}(m_o, \bar{v}_{k_o}, \ldots, \bar{v}_{N-2})$ (analogously V_L and V_N) then we have straightforwardly from Lemma 2, the following

Theorem 8

The regulator consisting of the feedback for complete state information of the state estimates produced by the Kalman filter given by (4.38-40) and (4.66-69) and with means V_R that solve

$$V_R = \min_{V_L \in \mathbb{R}^{n+r(N-1-k_o)}} (V^* - V_L)^T (JK)^T \sum_{k=k_o}^{N-1} \Phi_k^T N_k \Phi_k (JK)(V^* - V_L) \qquad (4.71)$$

$$V^* = \max_{V \in M} (V - V_R)^T (JK)^T \sum_{k=k_o}^{N-1} \Phi_k^T N_k \Phi_k (JK)(V - V_R)$$

$$+ V^T (IS) V - (V - V_N) D (V - V_N) \qquad (4.72)$$

is robust for the game $(M \times C, \mathcal{H}, J)$.

4.4 Continuous-time case

Consider a system described by the stochastic differential equations

$$dx_t = A(t)\, x_t\, dt + B(t)\, u_t\, dt + d\zeta_t \qquad (4.73)$$

$$dy_t = C(t)\, x_t\, dt + d\theta_t \qquad (4.74)$$

with x_t an n×1 state vector, u_t and m×1 control vector, y_t and r×1 output vector, and ζ_t, θ_t Wiener processes with incremental covariance

$$\text{cov}\, [d\zeta_t,\, d\theta_t] = \begin{bmatrix} \Xi(t) & \Psi(t) \\ \Psi^T(t) & \Theta(t) \end{bmatrix} dt \qquad (4.75)$$

and with the initial state $x(t_o)$ a gaussian random vector with mean x_o and variance Σ_o uncorrelated with ζ_t and θ_t.

We will suppose here that the means are known and that the covariance matrices belong to an uncertainty set $(\Xi(t), \Psi(t), \theta(t), \Sigma_o) = \Lambda \in C$. It is well known that the optimal observer of the state, i.e. the system that minimizes

$$L(H,\Lambda) = \int_{t_0}^{t_1} E[e_s^T Q(s) e_s | u_{<s}, y_{<s}]\, ds \tag{4.76}$$

where $e_t = x_t - \hat{x}_t$ and $Q(t)$ is a positive definite weighting matrix in the Kalman filter, is given by

$$d\hat{x}_t = A(t)\hat{x}_t dt + B(t)u_t dt + K(t)[dy_t - C(t)\hat{x}_t dt] \tag{4.77}$$

$$K(t) = [P(t)C^T(t) + \Psi(t)]\Theta^{-1}(t) \tag{4.78}$$

$$\begin{aligned} \dot{P}(t) = {} & [A(t) - \Psi(t)\Theta^{-1}(t)C(t)]P(t) + P(t)[A(t) - \Psi(t)\Theta^{-1}(t)C(t)]^T \\ & - P(t)C^T(t)\Theta^{-1}(t)C(t)P(t) + \Xi(t) - \Psi(t)\Theta^{-1}(t)\Psi^T(t) \end{aligned} \tag{4.79}$$

$$P(t_o) = \Sigma_o, \; x_{t_o} = x_o \tag{4.80}$$

Moreover, the optimal regulator for linear quadratic optimal control is the system that minimizes

$$J(H,\Lambda) = E[x_t^T F x_t | y_{<t}] + \int_{t_0}^{t_1} [x_s^T Q_1(s) x_s + u_s^T Q_2(s) u_s]\, ds \tag{4.81}$$

is given by

$$u(t) = -G(t)\hat{x}(t) \tag{4.82}$$

$$G(t) = -Q_2^{-1}(t)B^T(t)S(t) \tag{4.83}$$

$$-\dot{S}(t) = A^T(t)S(t) + S(t)A(t) + Q_1(t) - S(t)B(t)Q_2^{-1}(t)B(t)S(t) \tag{4.84}$$

$$S(t_1) = F \tag{4.85}$$

and $\hat{x}(t)$ given by (4.77-80).

Then, all the results in discrete-time have their respective parallels in continuous time:

Lemma 6

Suppose C' is a convex set and $(\Xi,\Psi,\Theta,\Sigma_o) = \Lambda = (1-\alpha)\Lambda_L + \alpha\Omega$ with $\Lambda_L = (\Xi_L,\Psi_L,\Theta_L,\Sigma_{oL})$ and $\Omega = (V,X,Y,Z)$ both belonging to C'.

Let $\bar{e}_t$ be the error of the optimal observer for Λ_L when Λ is the operating point and let $R_t = E[\bar{e}_t\bar{e}_t^T]$. Then

$$\lim_{\alpha \to 0} \frac{1}{\alpha} [P_t - R_t] = 0 \qquad (4.86)$$

Proof

By similar manipulations we can get the counterpart of equation (4.26):

$$(\dot{P} - \dot{R}) = (A - K_L C)(P - R) + (P - R)(A - K_L C)^T + (K - K_L)\Theta(K - K_L)^T \qquad (4.87)$$

Analogously to Appendix 1 we can see that $(K - K_L) = O(\alpha)$, and when α tends to 0, all the coefficients of the Ricati equation (4.79) are perturbed linearly. Moreover, since the initial condition $P(t_o) - R(t_o) = 0$, the solution of (4.87) is $o(\alpha)$.

Theorem 9

The Kalman filter for the least favorable collection of covariances for continuous-time state estimation is robust for the game $(C,\mathcal{K},L)$.

Proof

Using Lemma 6 the proof is completely analogous to the one in Theorem 1, noting that

$$L(H,\Lambda) = \int_{t_0}^{t_1} tr\{Q_t P_t\} dt \qquad (4.88)$$

For linear quadratic optimal control, the following Lemma [18] is used.

Lemma 7

$$E[X_{t_1}^T \; F \; X_{t_1} + \int_{t_0}^{t_1} [X_t^T \; Q_1 \; X_t + U_t^T \; Q_2 \; U_t] dt] =$$

$$= X_o^T \; S(t_o) X_o + tr\{S(t_o)\Sigma_o\} + \int_{t_0}^{t_1} tr\{\Xi S\} dt$$

$$+ E[\int_{t_0}^{t_1} (U_t + G \; X_t)^T \; Q_2 (U_t + G \; X_t) dt] \qquad (4.89)$$

where G and S are given by (4.83-85).

Theorem 10

The regulator consisting of the feedback for complete state information, of the state estimates produce' by the continuous-time Kalman filter for the least favorable collection of covariances for linear quadratic optimal control is robust for the game $(C,\mathcal{H},J)$.

Proof

Again, the penalty function is

$$J(H, \;) = X_o^T \; S(t_o) X_o + tr\{S(t_o)\Sigma_o\} + \int_{t_0}^{t_1} tr\{\Xi S\} dt$$

$$+ \int_{t_0}^{t_1} tr\{G^T \; Q_2 \; G \; P\} dt \qquad (4.90)$$

and Lemma 6 can be applied to yield the desired result by the same reasoning as in the proof of Theorem 6.

Identically to (4.31) we define in the continuous-time case:

$$W_\lambda(R) \triangleq \int_\lambda^{t_1} \Phi_L^T(t,\lambda)\, R_t\, \Phi_L(t,\lambda)\, dt \tag{4.91}$$

Then the next two theorems are proved similarly to Theorems 2 and 7.

<u>Theorem 11</u>

If the covariance uncertainty classes are independent then $(\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ is the least favorable collection of covariances for continuous-time state estimation if and only if the following are satisfied

$$\Xi_L = \arg \max_{\Xi \in X} \operatorname{tr} \left\{ \int_{t_1}^{t_0} \Xi(t) W_t(Q) dt \right\} \tag{4.92}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr} \left\{ \int_{t_1}^{t_0} \Psi(t) K_{Lt}^T W_t(Q) dt \right\} \tag{4.93}$$

$$\Theta_L = \arg \max_{\Theta \in N} \operatorname{tr} \left\{ \int_{t_1}^{t_0} \Theta(t) K_{Lt}^T W_t(q) K_{Lt} dt \right\} \tag{4.94}$$

$$\Sigma_{oL} = \arg \max_{\Sigma_o \in S} \operatorname{tr}\{\Sigma_o W_{t_o}(Q)\} \tag{4.95}$$

$$K_L = [P_L C^T + \Psi_L] \Theta_L^{-1} \tag{4.96}$$

$$\dot{P}_L = [A - \Psi_L \Theta_L^{-1} C] P_L + P_L [A - \Psi_L \Theta_L^{-1} C]^T - P_L C^T \Theta_L^{-1} C P_L + \Xi_L - \Psi_L \Theta_L^{-1} \Psi_L^T \tag{4.97}$$

$$P_L(t_o) = \Sigma_{oL} \tag{4.98}$$

Theorem 12

If the covariance uncertainty classes are independent then $(\Xi_L, \Psi_L, \Theta_L, \Sigma_{oL})$ is the least favorable collection of covariances for linear quadratic optimal control if and only if (4.96-98) and the following are satisfied.

$$\Xi_L = \arg \max_{\Xi \in X} \operatorname{tr} \left\{ \int_{t_0}^{t_1} \Xi(t)\, [S(t) + W_t(N)]\, dt \right\} \tag{4.99}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr} \left\{ \int_{t_0}^{t_1} \Psi(t) K_{Lt}^T W_t(N)\, dt \right\} \tag{4.100}$$

$$\Theta_L = \arg \max_{\Theta \in N} \operatorname{tr} \left\{ \int_{t_0}^{t_1} \Theta(t) K_{Lt}^T W_t(N) K_{Lt}\, dt \right\} \tag{4.101}$$

$$\Sigma_{oL} = \arg \max_{\Sigma \in S} \operatorname{tr}\{\Sigma_o (S(t_o) + W_{t_o}(N)\} \tag{4.102}$$

$$N = G^T Q_2 G \tag{4.103}$$

Steady state case

Consider the case in which the system and covariance matrices are time invariant. Under the proper assumptions, $S(t)$ and $P_L(t)$ given by the Ricatti equations (4.84) and (4.97) reach a steady-state solution. If we consider the case $t_1 = T \to \infty$, the penalty function should be converted to $\lim_{T \to \infty} \frac{1}{T} J(T)$, in which case, (4.99-101) result in

$$\Xi_L = \arg \max_{\Xi \in X} \operatorname{tr}\{\Xi(S + W(N))\} \tag{4.104}$$

$$\Psi_L = \arg \min_{\Psi \in P} \operatorname{tr}\{\Psi K_L^T W(N)\} \tag{4.105}$$

$$\Theta_L = \arg \max_{\Theta \in N} \operatorname{tr}\{K_L \Theta_L K_L^T W(N)\} \tag{4.106}$$

but, from (4.91) we can see that W(N) is the solution to the algebraic Riccati equation

$$(A - K_L C)^T W(N) + W(N)(A - K_L C) + N = 0 \qquad (4.107)$$

Therefore, (4.104) and (4.106) are precisely the equations given in [11] for the uncorrelated steady-state case. An analogous discussion with Theorem 11 results in the conditions of [10] for the state estimation problem.

4.5 Application of the Conditions for Least Favorability

The recursive character of the sets of equations for least favorability given in Sections 2 and 3 makes them attractive for a numerical solution based on an iterative search of least favorable covariances and filter matrices. However, for some uncertainty classes further analytical results are possible, as we show in this section. In order to illustrate the use of the conditions for least favorability we will consider several cases:

i) If the uncertainty classes do not impose a time dependence upon their elements (i.e., they constrain only the instantaneous values of the matrices), the least favorability conditions result in the extremization of the "weighted trace", $tr(U_k T_k^T)$, of the unknown matrix U for every sample, where T is given by every particular condition.

a) When the autocovariance classes χ, $\mathcal{N}$ or $\mathcal{S}$ have an element that is maximal for every time k, i.e. $U_{Mk} - U_k$ is positive semidefinite for every element U_k in the class, for all k, this is the least favorable (cf. [12],[14]). This follows because in the corresponding conditions every unknown autocovariance matrix appears multiplying, under the trace, a positive semidefinite matrix.

b) Suppose all is known of a particular covariance matrix U_k, is that its elements lie between some independent bounds

$$\underline{U}_k(i,j) < U_k(i,j) < \overline{U}_k(i,j) \qquad (4.108)$$

preserving the definiteness of the matrix if it belongs to an autocovariance class. Obviously this class is convex and the least-favorability conditions can be used. We get easily that the least favorable matrix satisfies

$$U_{Lk}(i,j) = \begin{cases} \overline{U}_k(i,j) & \text{if} \quad T_k(i,j) < 0 \\ \underline{U}_k(i,j) & \text{if} \quad T_k(i,j) > 0 \end{cases} \qquad (4.109)$$

In particular, if T_k is positive definite the diagonal elements of U_{Lk} have their largest possible value, a result that was achieved employing different techniques in [13] for observer design with uncertainties only in the autocovariances of the process and observation noises.

ii) Consider the following model for the deviation of the time varying matrices from a given nominal:

$$\mathcal{Y} = \left\{ \{U_k\}_{k=k_o}^{M},\ U_k \in \mathbb{R}^{a\times b}, \|U - U^{(o)}\| < \Delta \right\} \qquad (4.110)$$

with the possible preservation of definiteness, as before. The norm of (4.110) is the root mean square of the Euclidean norm, i.e.

$$\|U\|^2 = \sum_{k=k_o}^{M} \sum_{i=1}^{a} \sum_{j=1}^{b} (U_k(i,j))^2 \qquad (4.111)$$

and it is easy to check that coincides with the norm associated with the inner product

$$< A,B > = \mathrm{tr}\{ \sum_{k=k_o}^{M} A_k \, B_k^T \} \tag{4.112}$$

where A_k, $B_k \in \mathbb{R}^{a\times b}$.

Note that this definition of the norm for matrices of arbitrary frame allows the accomodation of uncertainties in the crosscovariance Ψ. Convexity of the set $\mathcal{Y}$ follows directly from the Schwarz inequality, thus it is possible to apply the conditions for least favorability, that can be expressed in general as

$$U_L = \arg \underset{U \in \mathcal{Y}}{\mathrm{ext}} < U,T > \tag{4.113}$$

where ext is max/min for auto/cross-covariances respectively. The solution of (4.113) is (Proposition 1, Chapter 3)

$$U_L = U^{(o)} \pm \varepsilon T \tag{4.114}$$

where ε is a nonnegative scalar such that U_L is an extremal element of $\mathcal{Y}$, and should be computed recursively. For example, suppose that the cross-covariance of the process and observation noise is uncertain around an uncorrelated nominal model. Then for the cases exposed in Sections 2 and 3, the least favorable conditions are (4.35), (4.42), (4.67), with the common form

$$\Psi_L = \arg \min_{\Psi \in \mathcal{P}} \mathrm{tr}\{ \sum_{j=k_o}^{N-2} \Psi_j \, K_{Lj}^T \, W_{j+1} \} \tag{4.115}$$

if the uncertainty class (4.110) is used, the least favorable is

$$\Psi_{Lj} = -\varepsilon\ W_{j+1}\ K_{Lj} \tag{4.116}$$

Substituting this into (4.38) the Kalman gain for the robust filter is

$$K_L = A\ \Sigma^L\ C^T\ [\,C\ \Sigma^L\ C^T + \Theta + \varepsilon\ WI]^{-1} \tag{4.117}$$

$$K_L = [\,(C\ \Sigma^L\ C^T + \Theta)\ I + \varepsilon\ W]^{-1}\ A\ \Sigma^L\ C^T \tag{4.118}$$

when the dimension of the system (n) and observation (r) is one respectively.

In order to treat uncertainties in the covariance of the initial condition, or when the other unknown matrices are restricted to be constant, we can use (4.110) as the deviation class with the additional assumption that U is time invariant and the norm (4.111) is not summed over time. Then we define the inner product as

$$\langle A,B \rangle = \mathrm{tr}\{A\ B^T\} \tag{4.119}$$

and the least favorable has the form

$$U_L = U^{(o)} \pm \varepsilon \sum_{k=k_o}^{M} T_k \tag{4.120}$$

if dealing with Ξ, Ψ or Θ, and (4.114) for Σ_o.

4.6 Conclusions

The application of the general formulation of minimax robust filtering has allowed us to present minimax theorems for predictor and filter state estimation and for quadratic control, under general classes of uncertainties in the second order statistics of the linear stochastic system. These results represent a generalization of earlier works [1],[2] devoted to the steady-state case for invariant continuous-time systems with uncorrelated noises. The minimax theorems state that the minimax filter is the optimal for the least favorable collection of noise covariances; this saddle point property implies a very attractive feature, namely, that when the actual covariances differ from the least favorable ones, the performance of the filter is upgraded. Sets of necessary and sufficient conditions are given for the least favorability of the four types of covariance matrices involved. These conditions lend themselves to a recursive solution, and are applied successfully to deviation classes of practical interest. It is worth to underscore that under the same type of uncertainty the worst case covariances do not necessarily coincide for the three types of filters treated here, in particular the observer used for the minimax regulator is not in general, a minimax predictor state estimator.

Given a particular expected deviation behavior, the tradeoff between the decrease of performance (with respect to the nominal filter) in the nominal model and the improvement of the worst case, should be assessed for the practical application of the minimax filters presented. The soft minimax approach has been used to deal with unknown means, and the connection with earlier works for estimators with unknown forcing functions, has been shown. If the covariance uncertainty classes can be modeled, as well, including nominals that are more likely to occur, their soft minimax solution is straightforward from the previous results.

Appendix 1

Proof of the existence of the linear term of the Taylor series of K in powers of α,

Denoting $\delta\pi = \lim_{\alpha \to o} \frac{1}{\alpha} (\pi(\alpha) - \pi(0))$, differentiating (3.17) and (3.18)

$$\delta K^* = ([A\,\delta\Sigma_k\,C^T + X] - [A\,\Sigma_L\,C^T + \Psi_L]R^{-1}[C\,\delta\Sigma_k\,C^T + Y])R^{-1} \qquad \text{(A.1.1)}$$

$$\begin{aligned}\delta\Sigma_{k+1} = {} & A\,\delta\Sigma_k\,A^T - [A\,\delta\Sigma_k\,C^T + X]\,R^{-1}\,[A\,\Sigma_L\,C^T + \Psi_L]^T \\ & + [A\,\Sigma_L\,C^T + \Psi_L]R^{-1}\,[C\,\delta\Sigma_k\,C^T + Y]\,R^{-1}\,[A\,\Sigma_L\,C^T + \Psi_L]^T \\ & - [A\,\Sigma_L\,C^T + \Psi_L]R^{-1}\,[C\,\delta\Sigma_k\,A^T + X^T] + \nabla \qquad \text{(A.1.2)}\end{aligned}$$

with $\delta\Sigma_{k_o} = Z$ and $R = C\,\Sigma_L\,C^T + \Theta_L$, but since Θ_L is positive definite and Σ_L is positive semidefinite, R is nonsingular. Thus δK exists for every time k, and all $\Lambda_L, \Omega \in C'$.

Appendix 2

For the case of uncorrelated and time invariant noises with a priori known covariance of initial state and first order statistics, a proof of the minimax theorem for Kalman gain design in continuous time, has appeared previously in the literature [19]. The authors show the concavity in the uncertainties and the convexity in the gain matrix, of the payoff function, and therefore are able to apply a standard game theoretic theorem. Unfortunately the payoff function is not convex in the Kalman gain as can be illustrated by a simple one-dimensional continuous time example:

$$\dot{x} = -x$$

$$y = x + w$$

where the initial state at $t = 0$ is known, and the variance of the white observation noise w is assumed to be Θ. The mean-square error of an observer with constant gain K, at time $t = \frac{1}{2}$ is easily obtained

$$E\,[\,e^2(\tfrac{1}{2})\,] = \frac{1}{2}\,\Theta\,(K^2/(K+1))(1-\exp(-K-1))$$

a function that is not convex in K. (Obviously the optimal K is zero, since no observations are needed in order to estimate the state without error.)

The mistake in the proof of the convexity, is committed by neglecting the dependence of the error covariance on the gain K when taking the derivative of a Riccati equation with respect to K.

5. OTHER APPLICATIONS

5.1 Wiener Filtering

Wiener filtering is one of the most developed applications of minimax robust filtering to date. The precursor work of Kassam and Lim [26] was followed by the formalization of Poor [27] and the study of the discrete-time causal Wiener filtering problem by Vastola and Poor [28]. The purpose of this section is to illustrate the application of the results of Chapter 2 to the noncausal uncorrelated case (the general situation can be handled in a similar fashion). While the application of the robust filtering theorem leads to a slight generalization (the signal and noise uncertainties are not required to be independent) of Theorem 1 of [27], the uncertainties separation theorem results in a set of conditions for least-favorability similar to those given in Chapters 3 and 4. (Note incidentally that in the area of robust Wiener filtering an atypical definition of least-favorability has been used. This confusion has its origin in an obscure discussion in [31].) Let us consider the penalty function

$$E(H;s,n) = \int_{-\infty}^{\infty} [s(\omega)|K(\omega) - H(\omega)|^2 + n(\omega)|H(\omega)|^2]\, d\omega \tag{5.1}$$

where $H \in \mathcal{H}$, $(s,n) \in P \subset F \times F$ with $\mathcal{H}$ and F representing some adequate spaces of filter transfer functions and power spectral densities, respectively.

Theorem 5.1

Suppose P is convex, then if (s_L,n_L) is a least favorable for $(P,\mathcal{H},E)$

$$H_L(\omega) = K_L(\omega)\,\frac{s_L(\omega)}{s_L(\omega) + n_L(\omega)} \tag{5.2}$$

is a minimax robust filter for the game $(P,\mathcal{H},E)$.

Proof

Note that the penalty function (5.1) is linear (and therefore concave) in the uncertainties. (Moreover, it is convex in the set of filters, thus standard minimax theorems could be used in this proof.) So it is sufficient to prove that $(H_L,(s_L,n_L))$ is a regular pair in order to prove the result via the robust filtering theorem.

Let $(\sigma,\nu) = (1-\alpha)(s_L,n_L) + \alpha(s,n)$, where $(s,n) \in P$. Then

$$E^*(\sigma,\nu) - E(H_L;\sigma,\nu) =$$

$$\int_{-\infty}^{\infty} |K(\omega)|^2 \left|\frac{\sigma(\omega)\nu(\omega)}{\sigma(\omega)+\nu(\omega)}\right| - \sigma(\omega)|K(\omega) - H_L(\omega)|^2 + \nu(\omega)|H_L(\omega)|^2 \, d\omega$$

$$= \int_{-\infty}^{\infty} |K(\omega)|^2 \left[\frac{\sigma(\omega)\nu(\omega)}{\sigma(\omega)+\nu(\omega)} - \sigma(\omega)\left[\frac{n_L(\omega)}{s_L(\omega)+n_L(\omega)}\right]^2 - \nu(\omega)\left[\frac{s_L(\omega)}{s_L(\omega)+n_L(\omega)}\right]^2\right] d\omega$$

$$= \int_{-\infty}^{\infty} - \frac{[\sigma(\omega)n_L(\omega) - \nu(\omega)s_L(\omega)]^2}{[n_L(\omega)+s_L(\omega)]^2(\sigma(\omega)+\nu(\omega))} |K(\omega)|^2 \, d\omega$$

$$= \int_{-\infty}^{\infty} - \alpha^2 \frac{[s(\omega)n_L(\omega) - n(\omega)s_L(\omega)]^2}{[n_L(\omega)+s_L(\omega)]^2(\sigma(\omega)+\nu(\omega))} |K(\omega)|^2 \, d\omega \qquad (5.3)$$

Hence, for every $(s,n) \in P$

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [E^*(\sigma,\nu) - E(H_L;\sigma,\nu)] = 0$$

Theorem 5.2

Suppose the uncertainties in signal and noise and independent, i.e. $P = S \times N$.

$((s_L,n_L),H_L)$ is a saddle point of $(S\times N,\mathcal{H},E)$ if and only if

1. $$H_L(\omega) = K(\omega) \frac{s_L(\omega)}{s_L(\omega) + n_L(\omega)} , \tag{5.2}$$

2. $$s_L = \underset{s \in S}{\operatorname{argmax}} \int_{-\infty}^{\infty} s(\omega) |K(\omega) - H_L(\omega)|^2 \, d\omega , \tag{5.4}$$

3. $$n_L = \underset{n \in N}{\operatorname{argmax}} \int_{-\infty}^{\infty} n(\omega) |H_L(\omega)|^2 \, d\omega . \tag{5.5}$$

Proof

The payoff function can be written as

$$-E(H;s,n) = f_1(H,s) + f_2(H,n) ,$$

$$f_1(H,s) = -\int_{-\infty}^{\infty} s(\omega) |K(\omega) - H(\omega)|^2 \, d\omega ,$$

$$f_2(H,n) = -\int_{-\infty}^{\infty} n(\omega) |H(\omega)|^2 \, d\omega ,$$

and therefore the uncertainties separation theorem (Thm. 2.2) can be used.

Note that (5.5) is a special case of (3.10) and therefore the previously derived results for matched filtering with noise uncertainty are useful here. Apropos of signal uncertainty the following result can be proved similarly to Prop. 3.1.

Proposition

The least favorable signal power spectral density in the class

$$\mathcal{S} = \left\{ s(\omega) \in \mathcal{F}: \int_{-\infty}^{\infty} |s_0(\omega) - s(\omega)|^2 \, d\omega \leq \Delta^2 \right\} \tag{5.6}$$

is given by

$$s_L(\omega) = s_0(\omega) + \delta\,|K(\omega) - H_L(\omega)|^2 \qquad (5.7)$$

where

$$\delta = \Delta / \left[\int_{-\infty}^{\infty} |K(\omega) - H_L(\omega)|^4 \, d\omega\right]^{1/2} \qquad (5.8)$$

At this point, the reader should not be confident on a theory that leads to a result such as the proposition above. Suppose that the power spectral density of the noise is fixed and is bounded away from zero. Let η be a constant such that

$$\sup_{\omega \in \mathbb{R}} \frac{s_0(\omega)}{s_0(\omega) + \delta\,|K(\omega) - H_L(\omega)|^2} < \eta < 1$$

Then, straightforward manipulations show that $\eta s_L \in S$. How is it possible that an attenuated version of s_L is more favorable than s_L itself? Indeed, the Wiener filter for s_L gives larger mean square error than the Wiener filter for ηs_L. The clue of the question is that we are not using the penalty function that we should; in practice it is more meaningful, in general, to minimize the MSE relative to the signal power than the MSE itself. Of course in the original Wiener filtering problem (with fixed input PSD's) both criteria coincide; however, in robust Wiener filtering a different penalty function should be used, e.g. the MSE relative to the input signal power,

$$L(H;s,n) = \frac{\int_{-\infty}^{\infty} [s(\omega)\,|K(\omega) - H(\omega)|^2 + n(\omega)\,|H(\omega)|^2]\, d\omega}{\int_{-\infty}^{\infty} s(\omega)\, d\omega} \qquad (5.9)$$

Note that when the signal uncertainty class is such that all its members have the same power (like those in [26] and [27]), this is equivalent to (5.1). However, none of the results developed in Chapter 2 can be applied to this function, on one hand it is not concave in the signal, and on the other it cannot be put in the form required by the uncertainties separation theorem. Therefore, robust Wiener filtering in this more general setting remains an interesting open problem.

5.2 Quadratic Receivers

The optimum quadratic receiver for the detection of a stochastic signal imbedded in additive noise, maximizes the deflection of the quadratic test statistic $\langle s+n, H(s+n)\rangle$; if the noise is gaussian, the deflection can be expressed as [23]

$$D(H;K_s,K_n) = \frac{tr^2\{H K_s\}}{tr\{HK_n HK_n\}} \tag{5.10}$$

where K_s and K_n are the signal and noise autocorrelation operators. The optimal filter is

$$H^*(K_s,K_n) = K_n^{-1} K_s K_n^{-1} \tag{5.11}$$

achieving a maximum deflection

$$D^*(K_s,K_n) = tr\{K_s K_n^{-1} K_s K_n^{-1}\} \tag{5.12}$$

Several generalized signal-to-noise ratios have been used besides the deflection [23]. In [24] it was given a generalization of (5.10) -- involving the variance of the test under the signal present hypothesis -- that allows an analytic solution for the optimal filter, namely

$$G(H;K_s,K_n) = \frac{tr^2\{HK_s\}}{tr\{\alpha_0^2\ HK_nHK_n + 2\alpha_0\alpha_1\ HK_nHK_s + \alpha_1^2\ HK_sHK_s\}} \tag{5.13}$$

$$H^*(K_s,K_n) = (\alpha_0 K_n + \alpha_1 K_s)^{-1}\ K_s(\alpha_0 K_n + \alpha_1 K_s)^{-1}$$

In order to apply the robust filtering theorem to these instances, we need to check the convexity of the payoff functions in the possible uncertainties. It is easy to see (by counterexample) that neither (5.10) nor (5.13) are convex in the noise operator for arbitrary K_s. And while (5.13) is not convex in K_s for arbitrary α_0, α_1, K_n, (it can be shown taking derivatives in the one-dimensional case), (5.10) is convex in K_s:

$$tr^2(H[(1-\alpha)K_{s1} + \alpha K_{s2}]) \leq (1-\alpha)tr^2(HK_{s1}) + \alpha tr^2(HK_{s2})$$

where the inequality follows from the linearity of the trace and the convexity of $f(x) = x^2$. Therefore the robust filtering theorem can only be applied to the maximum-deflection quadratic receiver for fixed noise autocorrelation.

Theorem 5.3

Suppose the signal autocorrelation uncertainty set $\mathcal{S}$ is convex, and K_L is the least favorable signal autocorrelation, then $H_L = K_n^{-1}\ K_L\ K_n^{-1}$ is a robust filter for $(\mathcal{K},\mathcal{S},D)$.

Proof

All we need to prove is that (H_L,K_L) is a regular pair. Let $K_s = (1-\alpha)K_L + \alpha K$, $K \in \mathcal{S}$

$$D^*(K_s,K_n) - D(H_L;K_s,K_n) =$$

$$= \frac{tr\{K_n^{-1} K_s K_n^{-1} K_s\} tr\{H_L K_L\} - tr^2\{H_L K_s\}}{tr\{H_L K_L\}}$$

$$= \frac{1}{tr\{H_L K_L\}} \Big[tr\{H_L K_L\}[(1-\alpha)^2 tr\{K_n^{-1} K_L K_n^{-1} K_L\}$$

$$+ 2\alpha(1-\alpha) tr\{K_n^{-1} K_L K_n^{-1} K\} + \alpha^2 tr\{K_n^{-1} K K_n^{-1} K\}]$$

$$- ((1-\alpha) tr\{H_L K_L\} + \alpha tr\{H_L K\})^2 \Big]$$

$$= \frac{\alpha^2}{tr\{H_L K_L\}} [tr\{K_n^{-1} K K_n^{-1} K\} tr\{H_L K_L\} - tr^2\{H_L K\}] \tag{5.14}$$

therefore (H_L,K_L) is indeed a regular pair for every $K_L \in S$.

When there are uncertainties in the noise autocorrelation the problem can still be solved via the following direct application of the uncertainties separation theorem.

Theorem 5.4

Suppose the uncertainties in the signal and noise autocorrelation operators are independent. $(H_L,(K_{sL},K_{nL}))$ is a saddle point of $(\mathcal{H},S\times N,D)$ if and only if the following conditions hold:

1. $$H_L = K_{nL}^{-1} K_{sL} K_{nL}^{-1} , \tag{5.15}$$

2. $$K_{sL} = \min_{K_s \in S} tr\{H_L K_s\} , \tag{5.16}$$

3. $$K_{nL} = \max_{K_n \in N} tr\{H_L K_n H_L K_n\} \tag{5.17}$$

Note that condition 2 has the same form that the least favorability conditions did in Chapter 4. In particular, the results in Section 4.5 hold for the signal uncertainty here with minor modifications. Note also that the uncertainties separation theorem cannot be applied to the generalized criterion (5.13), providing another example in which neither of the theorems in Chapter 2 can be used.

5.3 Output-Energy Filter

When the received signal is random, sometimes a linear receiver that maximizes the signal to noise ratio is used [25]. In this case the output signal-to-noise ratio is given by:

$$R(h;K_s,K_n) = \frac{\text{Energy of output due to signal}}{\text{Energy of output due to noise}}$$

$$= \frac{\langle h,\ K_s h\rangle}{\langle h,\ K_n h\rangle}\ , \tag{5.18}$$

where as before K_s and K_n are the signal and noise autocorrelation operators. Further justification of this criterion of optimality is possible considering the SNR of (3.1) as a random variable, in which case its expected value coincides with (5.18). Moreover, consider the special type of quadratic receiver consisting of a linear filter followed by a square law device. If the signal and noise are jointly gaussian processes, the variance of the output of the receiver is proportional to the square of the variance of the output of the filter, and therefore the deflection of the output of the receiver is proportional to the square of the energy ratio (5.18).

The optimal output-energy filter is

$$h^*(K_s,K_n) = \text{eigenvector of } K_n^{-1}K_s \text{ with maximum eigenvalue} \tag{5.19}$$

This can be proved analogously to Theorem X.13 of [29].

Claim

$R(h;K_s,K_n)$ is

1. Convex in K_s for arbitrary h and K_n;
2. Convex in K_n for arbitrary h and K_s;
3. Not convex in (K_s,K_n) for arbitrary h.

Proof

The first assertion is evident since $R(h,K_s,K_n)$ is linear in K_s. To prove the second we use the method of the proof of Lemma 6 of [30]. Let $K_n = (1-\alpha)K_1 + \alpha K_2$ and $\beta = \alpha \frac{\langle h, K_1 h\rangle}{\langle h, Kh\rangle}$. Then

$$\frac{1}{\langle h, K_n h\rangle} = \alpha \frac{1}{\langle h, K_n h\rangle} + (1-\alpha) \frac{1}{\langle h, K_n h\rangle}$$

$$= \beta \frac{1}{\langle h, K_1 h\rangle} + (1-\beta) \frac{1}{\langle h, K_2 h\rangle}$$

$$\leq \alpha \frac{1}{\langle h, K_1 h\rangle} + (1-\alpha) \frac{1}{\langle h, K_2 h\rangle} \tag{5.20}$$

where the inequality follows from $\beta \leq \alpha$ (K_1,K_2 are non-negative operators by assumption). The third assertion is easily verified b) a one-dimensional counterexample. (Let $(K_s,K_n) = \frac{1}{2}(4,2) + \frac{1}{2}(1,1)$.)

Therefore the robust filtering theorem, if applicable, will treat uncertainties either in the signal or noise operator but not in both. However, a general result of the kind that we have presented for all the previous filtering situations is not possible here, because the existence of regular pairs for given operators K_{sL} and K_{nL} will depend on every particular uncertainty class. As an example consider the simple discrete-time case in which $K_n = I$ and we have some class S of diagonal matrices.

Then any matrix with distinct diagonal elements forms a regular pair with its maximum-eigenvalue eigenvector; however, any matrix whose maximum-eigenvalue eigenspace has dimension greater than unity does not form a regular pair with any vector. In order to see this let $K_s = \mathrm{diag}\{s_1,\ldots,s_n\}$, $K_L = \mathrm{diag}\{\ell_1,\ldots,\ell_n\}$, $K = \mathrm{diag}\{k_1,\ldots,k_n\}$, $K_s = (1-\alpha)K_L + \alpha K$. Then (h_L,K_L) is a regular pair if and only if for every $K \in S$

$$h^T K_s h \,\|h_L\|^2 - h_L^T K_s h_L \,\|h\|^2 = o(\alpha) \tag{5.21}$$

where h is a maximum eigenvalue eigenvector of K.

If $\ell_i \neq \ell_j$ for all pairs there exists a positive β such that $0 \leq \alpha < \beta$ implies that $h = h_L$ for all $K \in S$, hence (5.21) is true. Nevertheless, if $\max(\ell_1,\ldots,\ell_n)$ is not unique (say $\max(\ell_1,\ldots,\ell_n) = \ell_1 = \ell_2$) there is no h_L for which (5.21) can hold for every $K \in S$. Of course (let $K = K_L$), in order for (5.21) to be true it is necessary that h_L is an eigenvector of K_L with maximum eigenvalue, i.e. $a(1,0,\ldots,0)$ and $b(0,1,0,\ldots,0)$, but if $k_1 < k_2$ and $h_L = a(1,0,\ldots,0)$ (analogously if $k_2 < k_1$ and $h_L = b(0,1,0,\ldots,0)$)

$$h^T K_s h \,\|h_L\|^2 - h_L^T K_s h_L \,\|h\|^2 = \|h\|^2 a^2 \alpha (k_2 - \ell_1) = O(\alpha) \tag{5.22}$$

Directly from the uncertainties separation theorem we have

<u>Theorem 5.5</u>

Suppose the uncertainties in the signal and noise autocorrelation operators are independent. Then $(h_L,(K_{sL},K_{nL}))$ is a saddle point of $(\mathcal{K}, S\times N, R)$ if and only if

1. h_L is an eigenvector of $K_{nL}^{-1}K_{sL}$ with maximum eigenvalue,

2. $$K_{sL} = \min_{K_s \in S} \langle h_L, K_s h_L \rangle , \tag{5.23}$$

3. $$K_{nL} = \max_{K_n \in N} \langle h_L, K_n h_L \rangle \tag{5.24}$$

Note that (5.24) is the same as condition 3 of Theorem 3.2, and that if there are minimal and maximal elements in S and N respectively, they are the solution of (5.23) and (5.24).

5.4 Hypothesis Testing and Estimation of Location

Our main goal here is to illustrate the application of the results previously derived for minimax robust filtering, to other problems; specifically in this last section we deal with the cases of robust hypothesis testing [31] and robust estimation of a location parameter [30], regarded as the classical starting points of minimax robust procedures in statistics and signal processing.

In the first problem, Huber [31] presents the robust solution for a particular uncertainty set of probability measures, namely a mixture class; however, to the author's knowledge, no result for general uncertainty classes is available yet. The application of the robust filtering theorem, although it imposes some restrictions on the possible uncertainty classes, in a step forward in that direction. On the other hand, in the robust estimation problem, Huber does provide a general theorem (T.2 of [30]) for convex uncertainty classes; here we will rederive that result via the robust filtering theorem. (It is worth noting that the game-theoretic results of robust filtering (e.g. [3],[27]) were in good part inspired by Huber's proof of (T.2 [30]).)

Suppose the problem is to decide with minimum probability of error (any other Bayes test can be treated similarly) between two probability measures upon observation of a random vector x. Hence the penalty function is

$$P_e(\phi, P_0, P_1) = \Pi_0 \int_\Omega \phi \, dP_0 + \Pi_1 \int_\Omega (1-\phi) dP_1 \tag{5.25}$$

where the test $\phi(x) \in \Phi$ is the conditional probability of deciding P_1 given that x is observed. The test that minimizes (5.25) is the likelihood ratio test:

$$\phi(x) = \begin{cases} 1 & \dfrac{dP_1}{dP_0} > \dfrac{\Pi_0}{\Pi_1} \\ \kappa & \dfrac{dP_1}{dP_0} = \dfrac{\Pi_0}{\Pi_1} \\ 0 & \dfrac{dP_1}{dP_0} < \dfrac{\Pi_0}{\Pi_1} \end{cases} \tag{5.26}$$

with arbitrary κ.

Now we assume that the pair of measures (P_0, P_1) is only known to belong to a certain set $\mathcal{Q}$. With respect to the game $(\mathcal{Q}, \Phi, P_e)$ we state and prove the following.

<u>Theorem 5.6</u>

Suppose that

1. The uncertainty set of probability measures $\mathcal{Q}$ is convex; and
2. for every pair $(M_0, M_1) \in \mathcal{Q}$,

$$\int I \left\{ \frac{dM_1}{dM_0} < \frac{\Pi_0}{\Pi_1} = \frac{dP_{1L}}{dP_{0L}} \right\} dx \int I \left\{ \frac{dP_{1L}}{dP_{0L}} = \frac{\Pi_0}{\Pi_1} < \frac{dM_1}{dM_0} \right\} dx = 0 \tag{5.27}$$

where I denotes the indicator function. Then, if (P_{0L},P_{1L}) is least favorable for $(\mathcal{Q},\Phi,P_e)$ it has an optimal test which is minimax robust.

Proof

Since (5.25) is linear in P_0 and P_1 it follows that it is concave in (P_0,P_1). Note that it is also concave in the priors, but it is not concave on (Π_0,P_0,Π_1,P_1). (Since it is also convex in Φ, the application of concave-convex minimax theorems could be investigated.) Therefore all we have to show is that for some test ϕ_L, $((P_{0L},P_{1L}),\phi_L)$ is a regular pair. To this end consider $(P_0,P_1) = (1-\alpha)(P_{0L},P_{1L}) + \alpha(M_0,M_1)$ for an arbitrary $(M_0,M_1) \in \mathcal{Q}$, then

$$P_e^*(P_0,P_1) - P_e(\phi_L,P_0,P_1) = \int_\Omega (\phi-\phi_L)(\Pi_0 dP_0 - \Pi_1 dP_1) \qquad (5.28)$$

with ϕ an optimal test for (P_0,P_1).

The region of integration of (5.28) will be divided into 9 regions:

$$\Lambda_1 = \left\{ x: \frac{dP_{1L}}{dP_{0L}} < \frac{\Pi_0}{\Pi_1} < \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_2 = \left\{ x: \frac{dP_{1L}}{dP_{0L}} = \frac{\Pi_0}{\Pi_1} < \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_3 = \left\{ x: \frac{dP_{1L}}{dP_{0L}} = \frac{\Pi_0}{\Pi_1} = \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_4 = \left\{ x: \frac{dP_{1L}}{dP_{0L}} < \frac{\Pi_0}{\Pi_1} = \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_5 = \left\{ x: \frac{dP_{1L}}{dP_{0L}} < \frac{\Pi_0}{\Pi_1} > \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_6 = \left\{ x:\ \frac{dP_{1L}}{dP_{0L}} > \frac{\Pi_0}{\Pi_1} < \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_7 = \left\{ x:\ \frac{dP_{1L}}{dP_{0L}} > \frac{\Pi_0}{\Pi_1} = \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_8 = \left\{ x:\ \frac{dP_{1L}}{dP_{0L}} = \frac{\Pi_0}{\Pi_1} > \frac{dP_1}{dP_0} \right\}$$

$$\Lambda_9 = \left\{ x:\ \frac{dP_{1L}}{dP_{0L}} > \frac{\Pi_0}{\Pi_1} > \frac{dP_1}{dP_0} \right\}$$

Thus, using a slight abuse of notation,

$$P_e^*(P_0,P_1) - P_e(\phi_L,P_0,P_1) =$$

$$\left[\int_{\Lambda_1} + \int_{\Lambda_2}(1-\kappa_L) + \int_{\Lambda_3}(\kappa-\kappa_L) + \int_{\Lambda_4}\kappa + \int_{\Lambda_7}(\kappa-1) - \int_{\Lambda_8}\kappa_L - \int_{\Lambda_9}\right](\Pi_0 dP_0 - \Pi_1 dP_1). \tag{5.29}$$

Now observe that we can write

$$\Pi_0 dP_0 - \Pi_1 dP_1 = \left(\frac{\Pi_0}{\Pi_1} - \frac{dP_{1L}}{dP_{0L}} + \frac{dP_{1L}}{dP_{0L}} - \frac{dP_1}{dP_0}\right)\Pi_1 dP_0$$

$$= \left(\frac{\Pi_0}{\Pi_1} - \frac{dP_{1L}}{dP_{0L}}\right) + \alpha\,\frac{dP_{1L}}{dP_{0L}} \cdot \frac{\Pi_1}{\Pi_0}\,(\Pi_0 dM_0 - \Pi_1 dM_1). \tag{5.30}$$

Therefore, taking into account the definition of the sets Λ_i,

$$\frac{1}{\alpha}\,[P_e^*(P_0,P_1) - P_e(\phi_L,P_0,P_1)] \geq$$

$$\geq \left[\int_{\Lambda_1} + \int_{\Lambda_2}(1-\kappa_L) + \int_{\Lambda_4}\kappa + \int_{\Lambda_7}(\kappa-1) - \int_{\Lambda_8}\kappa_L - \int_{\Lambda_9}\right]$$

$$\left(\frac{dP_{1L}}{dP_{0L}}\,\frac{\Pi_1}{\Pi_0}\,(\Pi_0 dM_0 - \Pi_1 dM_1)\right) \tag{5.31}$$

and

$$[\int_{\Lambda_2}(1-\kappa_L) - \int_{\Lambda_8}\kappa_L]\,(\frac{dP_{1L}}{dP_{0L}}\frac{\Pi_1}{\Pi_0}\,(\Pi_0 dM_0 - \Pi_1 dM_1) =$$

$$= [\int_{\Lambda_2}(1-\kappa_L) - \int_{\Lambda_8}\kappa_L]\,(\Pi_0 dM_0 - \Pi_1 dM_1), \tag{5.32}$$

which is always zero with an appropriate choice of κ_L, namely, $\kappa_L = 0$ or 1 if the first or second integral, respectively, of (5.27) is nonzero. If both integrals are zero the value of κ_L is irrelevant. Note that we have used the fact -- see eq. (5.30) -- that

$$\frac{dM_1}{dM_0} \geq \frac{\Pi_0}{\Pi_1} = \frac{dP_{1L}}{dP_{0L}} \quad \text{iff} \quad \frac{dP_1}{dP_0} \geq \frac{\Pi_0}{\Pi_1} = \frac{dP_{1L}}{dP_{0L}}$$

Since the RN derivatives are right continuous at $\alpha = 0$, $\Lambda_1, \Lambda_4, \Lambda_7, \Lambda_9$ reduce to null sets as $\alpha \downarrow 0$; therefore (5.31) and (5.32) result in

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha}\,[P_e^*(P_0,P_1) - P_e(\phi_L,P_0,P_1)] \geq 0 \ . \tag{5.33}$$

But for every α,

$$P_e^*(P_0,P_1) \leq P_e(\phi_L,P_0,P_1) \ , \tag{5.34}$$

therefore the regularity of $((P_{0L},P_{1L}),\phi_L)$ is demonstrated. It is interesting to underline that although the probability of deciding one of the hypotheses when the likelihood ratio is equal to the threshold is irrelevant for fixed probability measures, there are instances in which this is no longer the case for the robust test, as we have just seen.

In Huber's famous formulation of the problem of robust estimation of location [30], the statistician selects an influence curve $\psi \in \Psi$ in order to maximize the payoff

$$K(\psi,f) = \frac{(\int \psi f' dt)^2}{\int \psi^2 f dt} \tag{5.35}$$

where f is the marginal probability density of the (independent) errors in the location measurements. The maximum value of (5.35) is achieved with $\psi = -f'/f$ and is given by (via the Schwarz inequality)

$$K^*(f) = \int \frac{(f')^2}{f} dt \tag{5.36}$$

If the density function f lies in a convex set F, our purpose is to find the robust estimator for the game (Ψ, F, K). Theorem 2 of [30] states that the robust estimator is the optimal ψ_L for the least favorable density f_L. A proof of that theorem can be carried out by showing that (ψ_L, f_L) is a regular pair (Lemma 6 of [30] shows that K is convex in the uncertainties). Let $f = (1-\alpha)f_L + \alpha g$ with $g \in F$; then

$$\frac{1}{\alpha}[K^*(f) - K(\psi_L,f)] = \frac{1}{\alpha}\int \frac{(f')^2}{f} dt - \frac{(\int \psi_L f' dt)^2}{\alpha \int \psi_L^2 f dt}$$

$$= \frac{1}{\alpha \int \psi_L^2 f dt} \iint \frac{f'(s)\psi_L(t)}{f(s)} [f'(s)f(t)\psi_L(t) - f(s)f'(t)\psi_L(s)]\, dt ds \tag{5.37}$$

If $m(t,s) \triangleq (f_L(s)g'(t) + g(s)f_L'(t))\psi_L(s)$, the term in brackets is $\alpha[m(t,s) - m(s,t)] + o(\alpha)$. Hence, when $\alpha \downarrow 0$, (5.37) results in

$$\lim_{\alpha \downarrow 0} \frac{1}{\alpha} [K^*(f) - K(\psi_L, f)] =$$

$$\frac{1}{\int \psi_L^2 f_L dt} \iint \psi_L(s)\psi_L(t)[m(t,s) - m(s,t)]\,dtds = 0 \tag{5.38}$$

This demonstrates that (ψ_L, f_L) is a regular pair, and thus the problem solved by Huber fits the framework developed in Chapter 2.

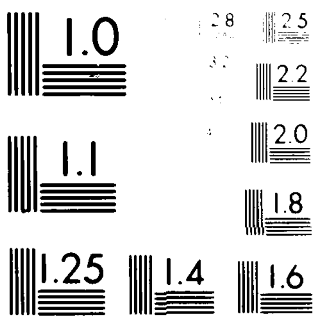
6. SUMMARY

A recapitulation of the main contributions of this work is given in this chapter.

First, the robust filtering theorem presented in Chapter 2 gives sufficient conditions on the performance function and on the sets of filters and operating points in order for the minimax robust filter to be the optimal filter for the least favorable operating point. In this theorem, the central issue turns out to be the introduction of the concept of regular pairs of filter and operating point. By means of this result, a systematic procedure is developed for solving problems in minimax robust filtering. So, in the light of this theorem, several previously treated problems have been revised and generalized, and other new problems have been proposed and solved. A technical refinement provided by the robust filtering theorem has been the introduction of an equivalent game by means of which typical restrictions of the uncertainty classes (pertaining to convexity and existence of least favorable elements) are relaxed in certain cases, and that allows the proof of extended results (class theorems) by considering restricted classes of filters. The next general result that we presented, i.e., the uncertainties separation theorem, constitutes a useful tool when dealing with independent (not necessarily convex) uncertainty classes. With some restrictions on the performance function, this theorem results in sets of equations that can be solved for the robust filter, numerically or analytically. Next, the concept of soft minimax has been introduced; it offers the possibility of incorporating in the model some possible further knowledge about the uncertain parameters by combining hard and soft constraints.

Following the presentation of these general results we have discussed their application to several specific filtering situations. First we have given a complete treatment of the robust matched filtering problem following the lines of the general work by Poor [3]; it is shown that the invertibility of the noise covariance operator is not necessary (but it is sufficient) in order for the regularity conditions to be satisfied, and a simple set of necessary and sufficient conditions for a pair of signal and noise to be least favorable is derived straightforwardly from the uncertainties separation theorem. Specific deviation classes are studied in the framework of finite-length discrete-time processing (although the majority of the results presented can be extended to the more general Hilbert space setting). Among other conclusions it is demonstrated that, when the deviations of the signal and of the noise covariance can be modeled by spheres centered on given nominals, the robust filter is the one matched to the nominal signal and to the nominal noise with an additional component of white noise (whose level is proportional to the size of the uncertainties); hence, in the frequent event in which the nominal noise is white (or in which the nominal signal is designed to be an eigenvector of the nominal noise covariance) the nominal matched filter is minimax robust. Concerning the optimal design of nominal signals for matched filtering in the presence of uncertainties, two extremal cases are examined, namely, the ℓ_2 uncertainty model, that results in a nonlinear optimization problem, and a particular case of the ℓ_∞ model, whose solution is the minimum-eigenvalue eigenvector of the noise covariance; i.e., the optimal signal without uncertainties.

The minimax state estimation and control problem is addressed next. General uncertainty classes of means and covariances of the process and


1.0
2.8
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

observation white noises are allowed, and the linear time-varying transient case is analyzed. Once again, the general results of Chapter 2 are applied to this case, leading to the desired necessary and sufficient conditions for least favorability, and to the demonstration of the minimax robustness of the optimal system for the least favorable statistics. In each of the problems that we consider, distinct conditions for least favorability are found; in particular, in relation to the separation theorem of stochastic control, it is shown that while the controller can be derived separately, its state estimates are not provided by the same Kalman filter as for minimax estimation. In order to deal with uncertain means, we employ the soft minimax approach introduced earlier; sufficient conditions for the nominal system to be (soft) minimax robust are provided. An interesting aspect is the fact that certain previous works in pursuit-evasion games with soft energy constraints are special cases of the problem solved here, and as such, arrive at coincidental solutions. Several examples of classes for uncertain noise auto and crosscovariances are studied in detail. Apart from their intrinsic interest they serve to illustrate the parallelism that we have encountered in dealing with analogous uncertainty classes in different filtering situations. This can be found, for instance, in the solution to the normed deviation model and in the transfer of uncertainties from one subclass to another already observed in the matched filtering problem.

Other filtering situations, such as the quadratic receiver and the output energy filter, in which the minimax robust approach has not been applied before are also investigated. In these cases, we point out several open problems for which the general results are not, as of now, applicable. For instance, no generalized performance function involving the variance of the output under the signal-present hypothesis leads to a solution of the robust quadratic receiver problem. Also, finding conditions for regularity in the output-energy filtering problem -- a generalization of the matched filter -- appears as a difficult and interesting question. The well-known problems of robust location estimation and of robust noncausal Wiener filtering fit the general framework of Chapter 2 and alternative proofs of their respective minimax robust theorems are provided. In the case of Wiener filtering, tractable and more meaningful approaches to deal with signal uncertainty classes without equal-power constraints should be investigated. Another interesting point is the study of the regularity condition when no explicit expressions for the optimal performance achievable at every operating point are available (e.g. causal filtering). Finally, another topic that deserves further attention is the minimax hypothesis testing problem for nonspecific uncertainty classes; although the regularity condition (5.27) required by our results does not seem to impose important restrictions in practice, its further relaxation appears to be interesting.

7. REFERENCES

[1] H. L. Royden, Real Analysis, 2nd Ed., McMillan Publ.: New York, 1968.

[2] V. Barbu, Th. Precupanu, Convexity and Optimization in Banach Spaces, Sijthoff & Noordhoff: Alphen Rijn, Holland, 1978.

[3] H. V. Poor, "Robust matched filters," IEEE Trans. Information Theory, to appear.

[4] G. L. Turin, "An introduction to digital matched filters," Proc. IEEE, vol. 64, no. 7, July 1976, pp. 1092-1112.

[5] S. Verdú, H. V. Poor, "Least favorable signals and noise for discrete time robust matched filtering," 1981 IEEE National Telecommunications Conference, Conf. Record, New Orleans, Louisiana.

[6] B. Noble, J. W. Daniel, Applied Linear Algebra, 2nd Ed., Prentice-Hall: Englewood Cliffs, NJ, 1977, pp. 431-434.

[7] H. Wielandt, Topics in the Analytic Theory of Matrices, Dept. of Mathematics, University of Wisconsin, Madison, 1976, pp. 26-27.

[8] A. V. Balakrishnan, Applied Functional Analysis, 2nd Ed., Springer-Verlag: New York, 1981.

[9] D. G. Luenberger, Optimization by Vector Space Methods, Wiley: New York, 1969.

[10] H. V. Poor and D. P. Looze, "Minimax state estimation for linear stochastic systems with noise uncertainty," IEEE Trans. Automatic Control, vol. AC-26, no. 4, Aug. 1981, pp. 902-906.

[11] D. P. Looze, H. V. Poor, K. S. Vastola and J. C. Darragh, "Minimax control of linear stochastic systems with noise uncertainty," IEEE Trans. Automatic Control, to appear.

[12] A. H. Jazwinski, Stochastic Processes and Filtering Theory, New York: Academic Press, 1970.

[13] J. O. Pearson, "Estimation of uncertain systems," in Control and Dynamic Systems, vol. 10, C. T. Leondes (ed.), New York: Academic Press, 1973.

[14] J. M. Morris, "The Kalman filter: A robust estimator for some classes of linear quadratic problems," IEEE Trans. Information Theory, vol. IT-22, no. 5, Sept. 1976, pp. 526-534.

[15] B.D.O. Anderson and J. B. Moore, Optimal Filtering, Englewood Cliffs, NJ: Prentice-Hall, 1979.

[16] H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, New York: Wiley-Interscience, 1972.

[17] B. Friedland, "On the effect of incorrect gain in Kalman filter," IEEE Trans. Automatic Control, vol. AC-12, no. 5, Oct. 1967, p. 610.

[18] K. J. Åström, Introduction to Stochastic Control Theory, New York: Academic Press, 1970.

[19] J. A. D'Appolito and C. E. Hutchinson, "A minimax approach to the design of low sensitivity state estimators," Automatica, vol. 18, no. 5, Sept. 1972, pp. 599-608.

[20] T. Basar and M. Mintz, "Minimax terminal state estimation for linear plants with unknown forcing functions," Int. Journal Control, vol. 16, no. 1, 1972, pp. 49-70.

[21] P. R. Halmos, Introduction to Hilbert Space, 2nd Ed., New York: Chelsea, 1957.

[22] C. R. Baker, "Optimum quadratic detection of a random vector in Gaussian noise," IEEE Trans. Communications, vol. COM-14, no. 6, Dec. 1966, pp. 802-805.

[23] W. A. Gardner, "Anomalous behavior of receiver output SNR as a predictor of signal detection performance exemplified for quadratic receivers and incoherent fading Gaussian channels," IEEE Trans. Information Theory, vol. IT-25, no. 6, Nov. 1979.

[24] S. Verdú, "Comments on 'Anomalous behavior of receiver output SNR as a predictor of signal detection performance exemplified for quadratic receivers and incoherent fading Gaussian channels,'" IEEE Trans. Information Theory, to appear.

[25] E. A. Robinson, Statistical Communication and Detection, New York: Hafner P.C., 1967.

[26] S. A. Kassam, T. L. Lim, "Robust Wiener filters," J. Franklin Institute, vol. 304, 1977, pp. 171-185.

[27] H. V. Poor, "On Robust Wiener filtering," IEEE Trans. Automatic Control, vol. AC-25, no. 3, June 1980, pp. 531-536.

[28] K. S. Vastola, H. V. Poor, "Robust Wiener-Kolmogorov theory," IEEE Trans. Information Theory, to appear.

[29] F. R. Gantmacher, The Theory of Matrices, Vol. 1, New York: Chelsea, 1960.

[30] P. J. Huber, "Robust estimation of a location parameter," Annals of Math. Statistics, vol. 35, 1964, pp. 73-101.

[31] P. J. Huber, "A robust version of the probability ratio test," Annals of Math. Statistics, vol. 36, 1965, pp. 1753-1758.

[32] D. Blackwell, M. A. Girshick, Theory of Games and Statistical Decisions, New York: Wiley, 1954.

[33] L. A. Zadeh, "On the definition of adaptivity," IEEE Proceedings, vol. 51, March 1963, pp. 469-470.

[34] S. A. Kassam, H. V. Poor, "A survey on robust filtering," IEEE 1982 International Symposium on Information Theory, Les Arcs, France, June 1982.

[35] M. C. Yovits, J. L. Jackson, "Linear filter optimization with game theory considerations," IRE Nat. Conv. Rec., 1955, pt. 4, pp. 193-199.

[36] N. J. Nilsson, "An application of the theory of games to radar reception problems," T.R. No. 704-1, 1958, Stanford Electronics Laboratories, Stanford, CA.

[37] L.-H. Zetterberg, "Signal detection under noise interference in a game situation," IRE Trans. Information Theory, vol. IT-8, 1962, pp. 47-57.

[38] A. A. Ershov, "Robust methods of evaluating parameters (survey)," Automation and Remote Control (U.S.S.R.), (English translation), no. 8, 1978, pp. 66-100.

[39] V. M. Krasnenker, "Stable (robust) detection methods for signal against a noise background (survey)," Automation and Remote Control (U.S.S.R.), (English translation), no. 5, 1980, pp. 640-659.

[40] S. Zahl, "The linear prediction of deterministic signals," IEEE Trans. Information Theory, vol. IT-10, no. 3, July 1964, pp. 222-226.

[41] D. E. Johansen, "Solution of a linear mean square estimation problem when process statistics are undefined," IEEE Trans. Automatic Control, vol. AC-11, no. 1, Jan. 1966, pp. 20-30.

[42] V. P. Kuznetsov, "Stable detection when the signal and spectrum of normal noise are inaccurately known," Telecomm. Radio Eng. (U.S.S.R.), (English translation), vol. 30-31, 1976, pp. 58-64.

[43] S. A. Kassam, T. L. Lim, L. J. Cimini, "Two dimensional filters for signal processing under modeling uncertainties," IEEE Trans. Geosci. Electron., vol. GE-18, 1980, pp. 331-336.

[44] D. R. Halverson, G. L. Wise, "An observation concerning signal to noise ratio properties of continuous and discrete time detectors," Proc. 19th Allerton Conference on Communication, Control and Computing, Univ. of Illinois, 1981, pp.

[45] C.-T. Chen, S. A. Kassam, "Robust multiple-input matched filters," Proc. 19th Allerton Conference on Communication, Control and Computing, Univ. of Illinois, 1981, pp.

[46] G. Moustakides, S. A. Kassam, "Robust Wiener filters for correlated signals and noise," Proc. 14th Conference on Information Sciences and Systems, Princeton University, 1980, pp. 308-313.

[47] L. J. Cimini, S. A. Kassam, "Robust and quantized Wiener filters for p-point spectral classes," Proc. 14th Conference on Information Sciences and Systems, Princeton University, 1980, pp. 314-319.

[48] H. V. Poor, "On binary communication through a distorting channel," Proc. 12th Southeastern Symposium on System Theory, Virginia Beach, May 1980.

[49] V. D. Vandelinde, R. Doraiswami, H. O. Yurtseven, "Robust filtering for linear systems," IEEE 1972 Conference on Decision and Control, New Orleans, Dec. 1972.

MED

8